

Enforcing Respect

Liberalism, Perfectionism, and Antidiscrimination Law

Matthew A. Shapiro
Balliol College, Oxford
DPhil in Politics
Trinity Term 2012

Abstract

Can contemporary liberalism justify antidiscrimination law? The question seems impertinent until we consider contemporary liberalism's commitment to limited government. Once we do, we realize that contemporary liberals may not complacently assume that their theories justify antidiscrimination law simply because discrimination based on race or sex is so obviously wrongful. Rather, they must scrutinize antidiscrimination law just as they do other regulation of individual conduct. Providing such scrutiny, this thesis argues that three of the most prominent contemporary liberal doctrines of political legitimacy—John Rawls's "political liberalism," an antiperfectionist version of the "harm principle," and Joseph Raz's "liberal perfectionism"—all fail to justify core applications of antidiscrimination law, applications that we intuitively consider perfectly legitimate. In light of this failure, contemporary liberalism faces a dilemma: it must jettison either its commitment to comprehensive, uniform antidiscrimination regimes or its antiperfectionism and overriding commitment to personal autonomy. This thesis argues for the latter course by providing an account of the wrongfulness of discrimination based on race or sex that condemns all instances of the conduct. According to this account, discrimination is wrong because acting on discriminatory intentions is wrong. More specifically, by taking another person's race or sex as a reason to treat her less favorably than one would treat people of other races or the other sex, one fails to respect her as a person, to regard her as a being of ultimate value. Unlike contemporary liberal accounts, this account is fully perfectionist, since it defines discrimination in terms of the intentions of discriminators, and the intentions of discriminators in terms of their attitudes, which partly constitute their moral characters. So long as we remain committed to antidiscrimination law in its current form, we must attend to discriminators' characters. And to attend to discriminators' characters, we must be willing to espouse perfectionism.

Enforcing Respect

Liberalism, Perfectionism, and Antidiscrimination Law

Matthew A. Shapiro
Balliol College

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Politics in the Department of Politics and International Relations at the University of Oxford.

Trinity Term 2012

Acknowledgments

Though nominally an independent undertaking, completing a doctoral thesis always turns out in reality to be a group project. Any claim of sole authorship is belied by the myriad forms of support the writer receives along the way. This thesis—and this writer—are no different.

Both of my supervisors, John Gardner and Dan Butt, read multiple drafts of each chapter and provided detailed and timely feedback. I am especially grateful to Professor Gardner for his patience and accessibility during the hectic last few months leading up to submission.

Dan McDermott was not officially one of my supervisors, but he might as well have been. His incisive “confirmation of status” report prompted me to rethink and rewrite large portions of the thesis, particularly Chapters 5 and 6. I blush to think of the (additional) errors that would have marred the thesis had it not been for his invaluable suggestions.

My examiners, Jeremy Waldron and George Sher, conducted a challenging but fair oral examination and offered penetrating and constructive comments, some of which I was able to respond to in this final version of the thesis.

For their numerous contributions, intellectual and personal, I wish to thank the friends I made at Oxford, particularly Stuart Baran, Richard Ekins, Will Feldman, Martin Luteran, Rachel Mazyck, Matt Pennycook, Steve Stromberg, and Paul Yowell. John William Devine deserves special mention, having spent countless hours discussing the substance of the thesis as well as enlivening the occasionally dreary existence of a graduate student with his good humor and captivating wit.

I completed the final stages of the writing process in the United States, while in law school. I simply could not have finished the thesis in such stressful circumstances without the support of Robert Leider and Fred Liu, both of whom spurred me to the finish line, offering heartfelt encouragement when it was sorely needed.

My parents, Mary and Howard Shapiro, managed to be as nurturing as they have always been while watching the thesis-writing process unfold from afar. But their contributions to this thesis run much deeper. From the beginning, they sought to instill in me an enduring intellectual curiosity. A doctorate is but a natural step along the path on which they set me long ago.

Typically, the most significant consequence of a doctoral thesis is the academic degree that signifies its completion. In my case, however, I came away with something far more important. It was only because I stayed on at Oxford to pursue a doctorate that I met my wife, Emily Stolzenberg. While Emily may have joined the project in medias res, it was she, more than anyone else, who bore the intellectual and emotional burdens of seeing it through. I dedicate this thesis to her.

Contents

Acknowledgments	ii
Introduction	1
Part I: Contemporary Liberalism and Antidiscrimination Law	
1. Liberal Legitimacy and Public Policy	29
2. Political Liberalism	58
3. The Harm Principle	115
4. Liberal Perfectionism	171
Part II: A Perfectionist Account of Antidiscrimination Law	
5. Discrimination and Respect	228
6. Respect and Politics	297
Conclusion	335
Works Cited	339

Introduction

I. CONTEMPORARY LIBERALISM AND ANTIDISCRIMINATION LAW

Can contemporary liberalism justify antidiscrimination law? The question seems impertinent until we consider contemporary liberalism's commitment to limited government, its insistence on restricting the exercise of political power in the name of individual liberty. Contemporary liberalism considers the paradigmatic form of individual liberty to be freedom of conscience. To secure this freedom, most contemporary liberal political theories espouse some version of antiperfectionism, the idea that the government must remain neutral among competing conceptions of the good and refrain from using political power to help its citizens to lead morally better lives. Antidiscrimination law seems to test the limits of contemporary liberalism's antiperfectionism, for we understand discrimination based on race or sex¹ not just as a harm or wrong suffered by its victims, but also as an immorality committed by its perpetrators. This thesis considers whether contemporary liberalism can justify the full extent of antidiscrimination law without having to jettison its antiperfectionism and, if not, what a perfectionist account of antidiscrimination law might look like.

There is no doubt that contemporary liberals themselves believe that their theories must justify antidiscrimination law in order to be compelling. Indeed, it seems that contemporary liberals would reject *as illiberal* any theory that condemned the legislation. Ronald Dworkin, for example, includes among the defining features of any liberal political theory a willingness to "support racial equality and approve government intervention to secure it, through constraints on both public and private discrimination in

¹ This thesis focuses exclusively on discrimination based on race or sex, which jurisdictions tend to regulate in common. It ignores discrimination based on other traits, such as religion, sexual orientation, and disability, which raise uniquely difficult questions.

education, housing, and employment.”² Similarly, John Rawls identifies our “conviction” that racial discrimination is unjust as one of our “considered judgments,” those “provisional fixed points which we presume any conception of justice must fit.”³ He then predicates the validity of liberalism itself on its ability to legitimate antidiscrimination law: “Justice as fairness, and other liberal conceptions like it,” he claims, “would certainly be seriously defective should they lack the resources to articulate the political values essential to justify the legal and social institutions needed to secure the equality of women and minorities.”⁴ Only if it can underwrite a rather robust antidiscrimination regime, Rawls and Dworkin suggest, should a liberal political theory command our allegiance.⁵

This perceived imperative to justify antidiscrimination law is matched by contemporary liberals’ complacency about the ability of their theories to satisfy that imperative. Such complacency characterizes Thomas Nagel’s divergent responses to discrimination and other sources of socioeconomic inequality. According to Nagel, we perceive a tension between our concern with political legitimacy and our commitment to socioeconomic equality. Nagel highlights the seriousness of this tension by considering attempts to eradicate two of the most incorrigible sources of socioeconomic inequality: class and talent. We balk at many equality-promoting measures, Nagel contends, because class divisions stem largely from the preference people bestow on their family members in general and their children in particular, a reasonable preference that merits a high degree of immunity from governmental interference. We thus must settle for policies that, as far as possible, consign familial bias to the personal sphere of the family but do not

² Ronald Dworkin, *A Matter of Principle* (Cambridge, Mass.: Harvard University Press, 1985), 187.

³ John Rawls, *A Theory of Justice*, rev. ed. (Cambridge, Mass.: Harvard University Press, 1999), 17-18.

⁴ John Rawls, *Justice as Fairness*, ed. Erin Kelly (Cambridge, Mass.: Harvard University Press, 2001), 66.

⁵ To be sure, nondiscrimination is not the only important theme in contemporary liberalism. Another theme is tolerance. See, e.g., T.M. Scanlon, *The Difficulty of Tolerance* (Cambridge: Cambridge University Press, 2003), ch. 10.

compromise the highly intimate relationships that obtain there.⁶ Because we cannot similarly relegate individual ambition to the personal sphere, we are even more wary of mitigating the contribution of talent to socioeconomic inequality, lest we license the government to interfere unduly with people's personal choices.⁷

Given Nagel's manifest anxiety about the legitimacy of public policies aimed at curtailing the inegalitarian effects of class and talent, we might expect him to display similar disquiet when he considers a third source of socioeconomic inequality: discrimination. But Nagel maintains his equanimity, dismissively noting that "[d]iscrimination is the product of a bad motive, prejudice, and there is nothing to be said for it."⁸ According to this view, we should respond to any conflict between prejudiced personal preferences and the imperative of equality not by attempting to reconcile the two, but rather by subordinating the former to the latter. Given his rightfully derisive condemnation of discriminatory motives and the sharp contrast he draws between class and talent, which raise concerns of political legitimacy, and discrimination, which does not, Nagel seems to assume that contemporary liberalism warrants a very extensive and intrusive antidiscrimination regime.

We should be careful, however, not to paint contemporary liberals with too broad a brush, for some liberal political philosophers seem to recoil from the most intrusive implications of their otherwise expansive conceptions of antidiscrimination law. Most notably, some liberals argue that religious associations like the Roman Catholic Church should enjoy significant immunity from antidiscrimination law—particularly prohibitions against sex discrimination. For example, Cass Sunstein would exempt religious associations from at least some provisions of antidiscrimination law, despite rejecting the "asymmetry thesis," which permits the application of ordinary civil and criminal law, but

⁶ Thomas Nagel, *Equality and Partiality* (Oxford: Oxford University Press, 1991), 109-12.

⁷ *Ibid.*, 112-19.

⁸ *Ibid.*, 109.

forbids the application of antidiscrimination law, to religious associations. In place of such an inflexible rule, Sunstein advocates a balancing test that weighs the importance of the interest underlying a particular provision of antidiscrimination law against the cost to a given religious association of governmental interference in its internal practices.⁹ Unlike the approach supported by the “asymmetry thesis,” Sunstein’s would subject religious associations to many provisions of antidiscrimination law, but it would nevertheless exempt them from some, an apparent departure from Nagel’s more expansive conception of antidiscrimination law.

Brian Barry envisions an even more limited antidiscrimination regime than does Sunstein. Indeed, he endorses the very “asymmetry thesis” Sunstein disavows, demanding a categorical exemption from prohibitions against sex discrimination for religious associations like the Catholic Church.¹⁰ Barry advocates immunizing even some secular private associations from antidiscrimination law, since he believes that not all discrimination violates equality of opportunity. While we should subject a particular private association to antidiscrimination law when necessary to guarantee equality of opportunity, Barry contends, we should otherwise permit the association to govern its internal practices as it chooses—even according to discriminatory principles and procedures.¹¹

Thus, while Nagel, Sunstein, and Barry all unhesitatingly condemn discrimination based on race or sex, they part company on the question of the *legitimacy* of various governmental remedies for discrimination, the appropriateness of particular applications of antidiscrimination law. Whereas Nagel takes the self-evident viciousness of discriminatory motives as sufficient warrant for the most extensive and intrusive

⁹ Cass R. Sunstein, “Should Sex Equality Law Apply to Religious Institutions?,” in *Is Multiculturalism Bad for Women?*, eds. Joshua Cohen, Matthew Howard, and Martha C. Nussbaum (Princeton: Princeton University Press, 1999), 86, 92-93.

¹⁰ Brian M. Barry, *Culture and Equality* (Cambridge: Polity Press, 2001), 174-76.

¹¹ *Ibid.*, 123.

antidiscrimination law, Sunstein and, to a greater extent, Barry insist that antidiscrimination law defer to the internal practices of certain kinds of associations. This latter position reflects the distinctively liberal concern with the propriety of governmental interference with individual conduct, a concern that ultimately requires the government to tolerate some wrongdoing that it would otherwise seek to eradicate. Evidently sharing this concern, Nagel condemns some governmental policies aimed at promoting socioeconomic equality. But unlike Sunstein and Barry, he considers questions of political legitimacy to be impertinent—if not obtuse—in the case of discrimination.

Which position better honors liberalism's fundamental tenets? Consider contemporary liberals' dichotomous responses to two other forms of wrongdoing: rape and adultery. No liberal would understand liberalism's concern with legitimacy to prevent the government from prohibiting every instance of rape. Echoing Nagel, because there is "nothing to be said for" rape or the motives that prompt it, the government need not respect those motives in devising the criminal law. But most liberals respond very differently to (at least most instances of) adultery, viewing it as the epitome of the kind of wrong that the government must tolerate. Given this distinction between two distinct categories of wrongdoing, one that raises concerns of legitimacy and another that does not, we might begin to question discrimination's status. Should we classify the discriminator's wrongdoing with the rapist's or the adulterer's? May we largely bracket questions of legitimacy when assessing antidiscrimination law, or must we confront them head-on?

Much turns on contemporary liberalism's response to such questions. For if the ideology does indeed subject antidiscrimination law to the strictures of political legitimacy, then contemporary liberals may no longer reflexively profess their allegiance to even the most central facets of antidiscrimination regimes, but rather must evaluate the regimes with the intense scrutiny they usually reserve for governmental responses to other

wrongs like adultery. And if this heightened inquiry reveals that the liberal concern with legitimacy actually condemns certain core applications of antidiscrimination law, then we will have exposed a profound inconsistency between fundamental liberal principles and public policies that we all—contemporary liberals included—consider to be especially important and undeniably legitimate.

The rest of this Introduction will show that contemporary liberal accounts of antidiscrimination law must indeed confront the question of political legitimacy, that contemporary liberals cannot assume that their theories justify antidiscrimination law simply because discrimination is so obviously wrongful. After providing a definition of “discrimination” and typology of antidiscrimination principles (Section II), this Introduction will expound the contemporary liberal concern with political legitimacy, particularly in its antiperfectionist guise (Section III), and then contrast that concern with perfectionist political theories (Section IV). This Introduction will conclude by suggesting that antidiscrimination law challenges contemporary liberal principles of political legitimacy and outlining the argument of the rest of this thesis (Section V).

II. DISCRIMINATION AND ANTIDISCRIMINATION PRINCIPLES

In attempting to define the concept of discrimination, a good place to begin is the actual antidiscrimination regimes of contemporary liberal democracies. For such regimes embody some of our most confident moral intuitions, identifying certain forms of conduct that we collectively deem sufficiently objectionable to warrant not just moral condemnation, but also legal prohibition. The United Kingdom’s antidiscrimination regime is typical. In particular, it identifies and prohibits two forms of discrimination: direct and indirect. According to U.K. antidiscrimination law, one *directly* discriminates against another person based on race or sex if, “because of” her race or sex, one treats her “less favourably than [one] treats or would treat other[.]” people of different races or the

other sex.¹² Again according to U.K. antidiscrimination law, one *indirectly* discriminates against another person based on race or sex if one “applies” to her a facially neutral “provision, criterion or practice” that one applies or would apply equally to people of other races or the other sex but that nevertheless disproportionately disadvantages members of her race or sex and for which one can offer no compelling race- or sex-neutral justification.¹³

Note that, so defined, both direct and indirect discrimination involve unfavorable or adverse treatment of a person because of her race or sex. At this point in the thesis, however, we should remain agnostic as to whether treating someone *more* favorably based on her race or sex also constitutes discrimination, for it may turn out that the best case against unfavorable race- and sex-based differential treatment also condemns at least some favorable race- and sex-based differential treatment. We should thus seek a definition of discrimination that comprehends both favorable and unfavorable race- and sex-based differential treatment while remaining faithful to the definitions embodied in contemporary antidiscrimination law. Kasper Lippert-Rasmussen provides such a definition. According to Lippert-Rasmussen,

X discriminates against (in favour of) Y in dimension W iff: (i) X treats Y differently from Z (or from how X would treat Z, were X to treat Z in some way) in dimension W; (ii) the differential treatment is (or is believed by X to be) disadvantageous (advantageous) to Y; and (iii) the differential treatment is suitably explained by Y’s and Z’s being (or believed by X to be) (members of) different, socially salient groups.¹⁴

This definition comprehends both direct and indirect discrimination: just as one satisfies all three elements if, “because of” a person’s race or sex, one treats her “less favourably than [one] treats or would treat others,” so one satisfies them if one applies to her a generally applicable policy (a form of “treatment”) that disproportionately affects

¹² Equality Act 2010, § 13(1).

¹³ *Ibid.*, § 19.

¹⁴ Kasper Lippert-Rasmussen, “The Badness of Discrimination,” *Ethical Theory and Moral Practice* 9 (2006): 168.

members of her sex or racial group (a form of “differential” treatment). The definition also explicitly comprehends both favorable and unfavorable differential treatment (without attempting to justify or condemn either kind of treatment). It is this definition that will guide this thesis’s analysis of various antiperfectionist accounts of antidiscrimination law and its attempt to develop its own, perfectionist one.

Although we should keep open the possibility that favorable race- and sex-based differential treatment is at least sometimes wrongful, this thesis will initially assess antiperfectionism’s ability to condemn only unfavorable differential treatment. This limited focus is justified because people are generally much more confident in their judgment that at least most forms of unfavorable race- and sex-based differential treatment are wrongful; the issue of favorable differential treatment, such as “affirmative action” or “positive discrimination,” is much more contentious. This thesis will thus attempt to construct the strongest possible account of the wrongfulness of unfavorable race- and sex-based differential treatment and only then consider what implications such an account might have for the issue of favorable differential treatment.

In addition, this thesis will follow Lippert-Rasmussen’s definition in treating direct and indirect discrimination as two instances of the same phenomenon. That is, it will posit a general category of “discrimination” and treat direct and indirect discrimination as both falling within it. It will thus assume that any account of the wrongfulness of discrimination (perhaps supplemented by other moral principles) should equally condemn both direct and indirect discrimination in order to be compelling. This assumption does not extend to the question of which form of discrimination is more fundamental: whereas some accounts initially condemn direct discrimination and then (again, perhaps relying on other moral principles) extend that condemnation to indirect discrimination, others proceed in the opposite direction. But either way, a compelling

account of discrimination, like contemporary antidiscrimination law itself, should attend to both forms of discrimination.

Despite recognizing two distinct forms of discrimination, contemporary antidiscrimination law typically enshrines only one kind of antidiscrimination principle. As Elisa Holmes has argued, antidiscrimination principles can be either deontic or telic. A deontic antidiscrimination principle simply prohibits certain individuals (employers, purveyors of public accommodations, etc.) from discriminating against others on certain grounds (race, sex, etc.); a telic antidiscrimination principle, by contrast, requires those individuals to perform certain actions designed to reduce the amount of discrimination suffered by certain protected groups.¹⁵ For example, a deontic antidiscrimination principle would simply instruct employers not to discriminate against any of their employees based on race, whereas a telic antidiscrimination principle would instruct them to make all employment decisions through committees with significant numbers of racial-minority members, thereby reducing the chances that the resulting decisions would be racially biased.¹⁶

Extant prohibitions against both direct and indirect discrimination are deontic, rather than telic, antidiscrimination principles. Just as the former prohibitions forbid employers and purveyors of public accommodations to treat individuals differently because of their race or sex, so the latter forbid them to treat individuals differently because of characteristics that closely correlate with race or sex (at least without adequate justification). Both, in other words, directly forbid discriminatory conduct. This thesis will thus assess antiperfectionist accounts of discrimination for their ability to justify the deontic principles enshrined in contemporary antidiscrimination law. Although telic

¹⁵ Elisa Holmes, "Anti-Discrimination Rights Without Equality," *Modern Law Review* 68 (2005): 182.

¹⁶ Cf. *ibid.*

antidiscrimination principles may have their virtues, contemporary antidiscrimination law has largely eschewed them in favor of deontic ones.

III. CONTEMPORARY LIBERALISM AND POLITICAL LEGITIMACY

To determine whether antidiscrimination law actually threatens to expose serious flaws in the contemporary liberal project, we must first examine more closely the project's theoretical underpinnings. According to Jeremy Waldron, liberalism founds itself on a distinctive conception of persons within their social environment. Demonstrating their strong affinity with the Enlightenment's vindication of reason, liberals conceive of persons as autonomous agents capable of subjecting social institutions to "critical scrutiny."¹⁷ Indeed, liberals maintain that "men and women reveal and exercise their highest powers as free agents when they engage in this sort of scrutiny of the arrangements under which they are to live."¹⁸ This general critical capacity of persons, in turn, entails a more specific ability "to determine for themselves how they will restrain their conduct in order to live in community with others" by proposing and endorsing the social rules according to which they will comport themselves.¹⁹

Liberals derive from their conception of the person a corresponding conception of public justification, a conception of how persons must justify social arrangements to their fellow citizens. More specifically, liberals insist that "all aspects of the social world should either be made acceptable or be capable of being made acceptable to every last individual."²⁰ In justifying public policies and institutions to one another, autonomous agents may invoke only reasons that everyone can reasonably accept as binding on his

¹⁷ Jeremy Waldron, *Liberal Rights* (Cambridge: Cambridge University Press, 1993), 41. Liberals are not always clear whether they are primarily concerned with "personal autonomy," being author of one's own life, or "moral autonomy," being author of the moral principles on which one acts, though they seem to focus on the former (see Jeremy Waldron, "Moral Autonomy and Personal Autonomy," in *Autonomy and the Challenges to Liberalism*, eds. John Christman and Joel Anderson (Cambridge: Cambridge University Press, 2005)).

¹⁸ Waldron, *Liberal Rights*, 41.

¹⁹ *Ibid.*, 42-43.

²⁰ *Ibid.*, 36-37.

own conscience; recondite reasons rooted in, say, religious dogma are inadmissible in public discourse. This discursive requirement ultimately proves rather stringent, given the diversity of beliefs and lifestyles in most contemporary western democracies, diversity that obscures “underlying interests and beliefs shared in common which may be appealed to in the justification of our institutional arrangements.”²¹ Nevertheless, liberalism establishes such shared interests and beliefs as the sole currency of public justification and, in so doing, distinguishes itself as an ideology.

Authority and the Nature of Political Power

In expounding liberalism’s conception of the person and concomitant conception of public justification, Waldron exposes some central liberal concerns, foremost among which is an intense suspicion of political power. After all, liberals would hardly trouble themselves to demand special justification for the exercise of political power if they did not somehow view it as a threat to their conception of the person as autonomous. But how, exactly, might political power compromise autonomy? Dworkin suggests a possible answer in articulating what he dubs the “classical problem of the legitimacy of coercive power,” which presents questions like the following: “How can *anything* provide even the general form of justification for coercion in ordinary politics? What can ever give anyone the kind of authorized power over another that politics supposes governors have over the governed?”²² These questions actually express two distinct normative concerns with political power—one with its coerciveness and another with its claim of authority—each of which might constitute the fundamental problem to which liberalism considers itself a response.

It seems natural for liberals to find the coerciveness of political power particularly unsettling, for that feature appears radically to contradict their conception of the person.

²¹ Ibid., 56.

²² Ronald Dworkin, *Law’s Empire* (Cambridge, Mass.: Harvard University Press, 1986), 191.

Roughly put, one individual, *A*, coerces another, *B*, when *A* compels *B* to act as *A* desires—either directly by using physical force against *B* or more obliquely by credibly threatening *B* with undesirable sanctions for recalcitrance. One obviously cannot function as a fully autonomous agent if a significant number of one’s choices result from coercion, so defined. And because political institutions employ coercion to enforce their dictates, it seems liberals must indeed demand a special justification for exercises of political power.

But this line of reasoning proceeds too quickly, exaggerating the extent to which political institutions rely on coercion and mistaking one of liberalism’s derivative normative concerns for its primary one. As Grant Lamond notes, we cannot attribute liberals’ discomfort with law to a more fundamental aversion to all coercion (though many liberals have such an aversion). For while many institutions in society other than political ones coerce people, liberals nonetheless concentrate on the threat posed to autonomy by the state in particular, a focus that the coerciveness of law alone cannot justify. In addition, as a matter of sociological and psychological fact, political institutions need not, and indeed do not, rely exclusively on coercion to render law efficacious. Rather, they also exploit Aristotelian “habituation,” social pressure, the allure of the benefits of social cooperation uniquely available under law, and other similar forces to ensure general compliance with legal directives. Coercion thus is, at most, one among many “prerequisites” for the existence of *effective* law, not law itself.²³ Most significant, though we might be tempted to identify law’s claim of *the right* to wield coercion and to regulate the wider use of coercion in society as the ultimate source of liberal unease about law, this feature turns out, on closer inspection, to be but one manifestation of law’s comprehensive claim of the right to regulate all matters arising in

²³ Grant Lamond, “Coercion and the Nature of Law,” *Legal Theory* 7 (2001): 45-50.

its jurisdiction.²⁴ It is the latter, more fundamental claim to which Dworkin's second question refers and that seems most directly to challenge basic liberal commitments.

Our brief exploration of the relationship between law and coercion thus has led us to the feature of political institutions potentially most threatening to liberalism: their claim of authority. This claim constitutes the core of what Leslie Green dubs the "self-image of the state," how political institutions and officials portray their relationships with their subjects. More specifically, the state presents itself as possessing the right to guide its subjects' behavior by issuing directives that, by the mere fact of their promulgation and regardless of the particular actions they enjoin, both give subjects a reason to comply and exclude from consideration some reasons for disobeying.²⁵ To be sure, other social institutions, such as the family, the church, and the military, claim an analogous right within their respective jurisdictions, but the state's claim of authority differs from theirs in at least two respects. First, the state represents its authority as "supreme": while other institutions in society might claim authority, the state denies the legitimacy of these claims except insofar as they answer to its own authority. Second, the state's authority is extensive in scope, purporting to bind all the people in a given territory regardless of their other allegiances.²⁶ The political institutions that constitute the state are unique among other social institutions, and political power is unique among other forms of social power, in claiming supreme and comprehensive authority.

Why might we think that the state's claim of authority challenges some of liberalism's fundamental principles? One widely alleged contradiction ultimately proves illusory. Robert Paul Wolff pronounces authority and autonomy incompatible because, put simply, the former entails "the right to rule" and the latter "the refusal to be ruled." As persons, we must decide for ourselves what reason requires of us, all things considered,

²⁴ Ibid., 54-56.

²⁵ Leslie Green, *The Authority of the State* (Oxford: Oxford University Press, 1988), 86.

²⁶ Ibid., 82, 84.

and must refuse to abdicate our responsibility for our actions to any other individual or to any institution—including the state. But, Wolff maintains, by recognizing the state’s authority, its purported right to supplant its subjects’ judgments with its own, one necessarily neglects one’s moral obligation to decide for oneself how best to act; hence the contradiction between autonomy and the state’s authority.²⁷

Wolff’s argument, however, fails to impugn the compatibility of authority and autonomy, for it miscomprehends the nature of authority, of autonomy, and of reason more generally. In particular, Wolff ignores the distinction between first- and second-order reasons. According to Joseph Raz, a first-order reason is a reason either to perform or to forgo an action, whereas a second-order reason is a reason either to act for a reason (a “positive second-order reason”) or to forgo acting for a reason (a “negative second-order reason” or “exclusionary reason”). What Raz terms a “protected reason” incorporates both a first-order reason to perform an action and an exclusionary reason to disregard certain reasons against so acting.²⁸ Authority, Raz contends, consists in the ability to issue protected reasons.²⁹ Wolff’s failure to apprehend the different species of reasons and, consequently, the true nature of authority leads him illicitly to equate the “principle of autonomy,” according to which “reason never justifies abandoning one’s autonomy, that is, one’s right and duty to act on one’s judgment of what ought to be done, *all things considered*,” with the “denial of authority,” the “false principle that there are no valid exclusionary reasons, that is, that one is never justified in not doing what ought to be done on the balance of first-order reasons.”³⁰ The two positions are, in fact, distinct, for reason can indeed require one to heed an exclusionary reason to act against the balance of first-order reasons. One can thus recognize authority without compromising

²⁷ Robert Paul Wolff, *In Defense of Anarchism* (Berkeley, Calif.: University of California Press, 1998), 18.

²⁸ Joseph Raz, *The Authority of Law* (Oxford: Oxford University Press, 1979), 17, 18.

²⁹ *Ibid.*, 29.

³⁰ *Ibid.*, 27.

one's autonomy, judging that one should, all things considered, disregard the balance of first-order reasons in favor of the decisions of an authoritative person or institution.

But while Wolff fails to demonstrate a strict inconsistency between authority and autonomy, Raz's typology of reasons and account of authority nevertheless indicate one important respect in which the two can at least be in tension with each other and thus why liberals might be wary of political power. When a person decides—however autonomously—to defer to an authority, he effectively surrenders his judgment over a certain range of issues, accepting the authority's determinations about how he should act as his own. In particular, by acknowledging the authority of his state, a citizen views its commands as binding on him and as preempting his own conclusions about reason's requirements. To second-guess the state and to weigh for oneself the first-order reasons for and against an action the state enjoins would be to controvert the state's claim of authority. Though not necessarily incompatible with autonomy, this implication of acknowledging the state's authority underscores the need for intense vigilance on the part of the autonomous person. For in surrendering his judgment to the state, he must be relatively confident that the state will not abuse its authority. He must therefore scrutinize the various institutions through which the state exercises its authority before acceding to its rule in order to assure himself of the existence of sufficient safeguards against governmental malfeasance. Because liberalism conceives of the person as autonomous, it will be particularly skeptical of the state's pretensions to authority, demanding a rather probing interrogation of political institutions and power.

Three Liberal Concerns with Political Authority

We can now begin to comprehend Waldron's account of liberalism's fundamental principles—in particular, his portrayal of liberalism as, “at bottom, a theory about what makes political action—and in particular the enforcement and maintenance of a social and

political order—morally legitimate.”³¹ As we have seen, political action is authoritative action: states claim authority and expect their subjects to accept this claim. When viewed through a conception of the person as autonomous, this feature prompts liberalism to demand special justification for political action, to oblige political institutions to defend their exercise of power as “morally legitimate.” But the notion of the “moral legitimacy” of political institutions and power actually refers to at least three distinct normative concerns, each of which has preoccupied liberals to varying degrees.

First and most straightforwardly, we might wonder why we need political institutions at all. Why, in ordering their interactions with one another, must human beings establish institutions that claim supreme, comprehensive authority over their subjects? To attempt to answer such questions is to undertake what A. John Simmons defines as the task of “justification” or “justifying the state,” of “showing that some realizable type of state is on balance morally permissible (or ideal) and that it is rationally preferable to all feasible nonstate alternatives.”³² Because the compatibility of political authority and autonomy ultimately depends on justifying the state, liberalism attaches significant importance to that task.

We must not only defend the existence of the state in the abstract, but also demonstrate how any individual state can acquire the authority it claims over its particular subjects. In Simmons’s terminology, we must prove the “legitimacy” of each state, “the complex moral right it possesses to be the exclusive imposer of binding duties on its subjects, to have its subjects comply with these duties, and to use coercion to enforce the duties.”³³ To be fair, Simmons does not expressly reduce “the complex moral right” in which a state’s legitimacy consists to the notion of political authority expounded above,

³¹ Waldron, *Liberal Rights*, 50.

³² A. John Simmons, “Justification and Legitimacy,” *Ethics* 109 (1999): 742.

³³ *Ibid.*, 746.

but the two are nevertheless connected, with the latter encompassing all the constituents of the former.

Allen Buchanan offers a more restricted account of political legitimacy, according to which “an entity has *political legitimacy* if and only if it is morally justified in wielding political power, where to wield political power is to (make a credible) attempt to exercise supremacy, within a jurisdiction, in the making, application, and enforcement of laws.”³⁴ Buchanan then depicts this definition as bracketing the questions of political obligation³⁵ and of what he terms “authoritativeness,” whether “the fact that [an entity] issues a rule can in itself constitute a compelling reason to comply with that rule.”³⁶ But Buchanan’s protestations to the contrary notwithstanding, the concepts of political legitimacy and authority are inextricably linked. For even Buchanan defines political legitimacy in terms of law, and, as Raz notes, “it is an essential feature of law that it claims legitimate authority.”³⁷ To pronounce a state politically legitimate, then, we must first assess its suitability for law, which, in turn, requires rendering a verdict on its authority. Liberalism in particular insists that individual states justify their claims of authority by, inter alia, demonstrating the rectitude of their institutions before subjecting autonomous persons to their dominions.

Both Simmons’s and Buchanan’s definitions of political legitimacy predicate that concept of states or political institutions, but other political philosophers also apply it to particular governmental actions, indicating a third problem surrounding political authority to which liberalism responds. Thomas Nagel, for example, identifies liberalism primarily with its distinctive account of the legitimacy not only of political “arrangements” and

³⁴ Allen E. Buchanan, *Justice, Legitimacy, and Self-determination* (Oxford: Oxford University Press, 2004), 235.

³⁵ *Ibid.*, 237-39.

³⁶ Allen E. Buchanan, “Political Legitimacy and Democracy,” *Ethics* 112 (2002): 692-96.

³⁷ Raz, *Authority of Law*, 30.

“institutions,” but also of the “requirements” they impose.³⁸ This latter application of the concept of political legitimacy concerns the proper limits of political authority rather than the grounds for its justification. Though a state’s authority might be supreme and extensive, it is not thereby unlimited or absolute, for there are some commands no authority may issue and some matters no distinctively *political* authority may address. The concept of political legitimacy thus also contemplates the task of distinguishing appropriate applications of governmental power from inappropriate ones.

Of the three normative concerns with political authority, the one that has most vexed contemporary liberal political philosophers, and the one most relevant to determining liberalism’s stance toward antidiscrimination law, is the third, that of political legitimacy in the sense of the limits of political authority and the propriety of particular applications of governmental power. Liberals focus on this problem because they perceive a tension between the principal purpose of political authority and the nature of modern moral culture. Liberals have long viewed political authority as a necessary means of adjudicating disputes about the requirements of morality.³⁹ Even when persons unite around the same general moral principles, they inevitably interpret the principles differently, contesting their application in particular circumstances. Only the state can resolve these disagreements authoritatively and, in so doing, secure the preconditions for peaceful, stable communal life.

In light of the profound moral divisions that plague modern political communities, however, liberalism has come to view this admittedly essential function of political authority with increasing ambivalence. It seems that moral disputes now extend beyond the question of how to apply general moral principles in particular circumstances to the question of which principles to recognize as valid in the first place. We contest, in other

³⁸ Thomas Nagel, “Moral Conflict and Political Legitimacy,” *Philosophy & Public Affairs* 16 (1987): 221.

³⁹ See, e.g., John Locke, *Two Treatises of Government*, ed. Peter Laslett (Cambridge: Cambridge University Press, 1998), II.9.

words, not only the resolution of various practical moral problems, but also the very terms of moral discourse. Faced with this radical moral disagreement, the state risks alienating its subjects through its authoritative dispositions of particular disputes. For if the disputants cannot agree even on principles to govern the adjudication itself, the unsuccessful party will likely reject any verdict as illegitimate, as presupposing basic normative commitments that he considers anathema and thus forfeiting its claim to his respect and acquiescence. One cannot simultaneously view the state's decisions in this way and recognize its claim of authority.

To reconcile autonomy and authority in the face of intractable disagreement about basic moral principles, contemporary liberalism seeks to ensure that the state exercises its authority only on terms all its subjects can countenance. More specifically, liberalism attempts to identify

principles to deal with conflict that can at some level be endorsed by everyone—principles that will both motivate and command respect and that will therefore give authority to results which are reached in accordance with them, even if those results do not in themselves command unanimous support.⁴⁰

If the state proceeds according to norms beyond those embodied in these principles, significant segments of its citizenry will flout its directives and reject its claim of authority. And yet, to found the state, its institutions, and its power on shared principles alone is effectively to limit political authority. For such principles compose but a portion—perhaps a very small one—of the larger reservoir of moral judgments on which each of us would draw to justify individual governmental decisions. They consequently license the government to rectify not every moral wrong we as individual persons would decry, but only a more limited class of immoral conduct. Contemporary liberalism accepts this perhaps regrettable implication of limiting political authority as a cost it must

⁴⁰ Nagel, *Equality and Partiality*, 23-24.

bear to honor its conception of the person as autonomous while permitting a role for the state as authoritative arbiter of moral disputes.

Liberal Doctrines of Political Legitimacy

All versions of contemporary liberalism recognize the need to limit political authority and to license the state to address only a restricted class of moral wrongdoing, but they nevertheless employ different methods to achieve that aim. More specifically, different versions of liberalism espouse different doctrines of political legitimacy, which specify the limited conditions under which the government may legitimately employ its power to regulate its citizens' behavior. Unless a proposed regulation satisfies these criteria, the government must stay its hand and abide at least some moral wrongdoing.

A doctrine of political legitimacy can circumscribe political authority and identify the class of moral wrongdoing amenable to governmental regulation in a number of ways. *Ends-limiting doctrines* constrain the government by limiting the ends it may pursue. More specifically, *direct ends-limiting doctrines* authorize the government to pursue only an explicitly specified end (say, harm prevention). *Indirect ends-limiting doctrines* function more obliquely, enumerating the normative criteria that the political institutions of a given political community must satisfy if any exercise of governmental power in that community is to be legitimate. Such doctrines seek in the first instance to regulate not individual exercises of governmental power, but rather the background conditions under which those exercises occur. They thus indirectly limit the ends the government may pursue, precluding laws that contravene the normative principles undergirding political institutions and authority.

Rather than dictating the grounds on which the government may employ its power, doctrines of political legitimacy might instead restrict the particular means by which the government may regulate its citizens' behavior. Unlike ends-limiting doctrines, such

means-limiting doctrines leave the government free to pursue whatever morally sound ends it deems important, simply specifying the various tools (the criminal code, civil law, the tax code, and so on) it may employ to realize those ends. Because they impose such restrictions in the name of personal autonomy, however, the doctrines nonetheless qualify as liberal.

While the different kinds of doctrines of political legitimacy thus deploy different government-limiting tactics, at least ends-limiting doctrines serve a common strategic purpose, a purpose that, as we have seen, contemporary liberalism takes as its *raison d'être*. For ends-limiting doctrines require the government to justify its exercises of political power in terms each of its citizens can accept as an autonomous agent. In so doing, they express liberalism's demand that we invoke political power not on the basis of moral principles we as individual persons happen to espouse, but only "on more restricted grounds—grounds which belong in some sense to a common or public domain."⁴¹ Were some citizens to flout this demand and to enlist political power to advance more partisan principles, they would present their compatriots with a dilemma: either controvert the state's claim of authority or accede to that claim and disavow some of their own fundamental moral beliefs. To forestall this dilemma, contemporary liberalism establishes a shared realm of public justification in which all exercises of political power must be grounded, a realm whose borders are demarcated by the liberal doctrines of political legitimacy. More specifically, the doctrines prohibit the government from regulating certain forms of behavior by requiring it to identify each wrong it proposes to rectify in the political community's public moral lexicon rather than the normative dialect of any particular group of citizens.⁴² According to contemporary

⁴¹ Ibid., 158. This demand echoes that of Rawlsian public reason. See John Rawls, *Political Liberalism*, paperback ed. (New York: Columbia University Press, 1996), 212-54.

⁴² Cf. Nagel, "Moral Conflict," 227-28 (claiming that liberalism seeks "a way of being impartial not only in the allocation of benefits or harms but in their identification").

liberalism, the government may remedy only those wrongs that can be identified and condemned in the terms of the correct liberal doctrine of political legitimacy.

IV. PERFECTIONISM AND POLITICAL LEGITIMACY

Understood as a political ideology, perfectionism rejects the liberal conception of political legitimacy, permitting the government to pursue almost all morally sound goals—even those that defy justification according to the terms of public reason. Unlike most contemporary liberal political theories, all perfectionist ones explicitly found themselves on conceptions of the good life, accounts of the valuable activities and relationships in which human flourishing consists. While such conceptions differ from one another considerably, deeming different relationships and activities valuable, most of them identify human flourishing with “some development of capacities or some achievement of excellence.”⁴³ Their respective judgments about the nature of the good life, moreover, all purport to be objective, independent of people’s various beliefs, wishes, and desires.⁴⁴ Relatively unimpressed by the existence of pervasive moral disagreement, perfectionism settles the controversial questions that contemporary liberalism seeks to bracket.

In endorsing a particular conception of the good life as objectively and uniquely valid, perfectionism need not commit itself to a fundamentalist, authoritarian political agenda.⁴⁵ For to determine the political program of a particular version of perfectionism, we must consider not only the version’s avowed ideological affiliation, but also its intrinsic philosophical content, the precise conception of the good it espouses. Indeed, as we shall discover in Chapter 4, a version of perfectionism might even adopt a conception of the good steeped in liberalism’s conception of the person, valuing personal autonomy

⁴³ Thomas Hurka, *Perfectionism* (Oxford: Oxford University Press, 1996), 4.

⁴⁴ *Ibid.*, 5.

⁴⁵ See, e.g., Joseph Chan, “Legitimacy, Unanimity, and Perfectionism,” *Philosophy & Public Affairs* 29 (2000): 10-18; Richard Arneson, “Perfectionism and Politics,” *Ethics* 111 (2000): 37-63.

as a constituent of the good life—a normative commitment that would forestall many of the portentous policies that contemporary liberals tend to impute to perfectionism. And though they might ascribe merely instrumental value to autonomy, other versions of perfectionism can nonetheless respect human diversity and constrain political power. Unless it is yoked to a particularly austere and uncompromising conception of the good life, perfectionism need not be oppressive.

As this attempt to debunk some of the misconceptions surrounding perfectionism suggests, we must be careful to distinguish between “philosophical” perfectionism, on the one hand, and “political” or “state” perfectionism, on the other.⁴⁶ For to endorse a conception of the good life is not necessarily to license the government to promote it through the law. Indeed, even some contemporary liberals claim to ground their respective political theories in (comparatively thin) conceptions of the good, from which they derive the strictures of liberal public justification and prohibitions against moralistic governmental regulation.⁴⁷ It is therefore in political perfectionism, a policy program of fostering human flourishing through the law, that contemporary liberalism confronts a true ideological adversary.

According to political perfectionism, the government may justify its policies not only in the narrow terms prescribed by liberal public justification, but according to any sound moral reasons that happen to support them.⁴⁸ Among these sound moral reasons, moreover, political perfectionism recognizes ones permitting the government to foster human flourishing directly by inculcating virtue and curbing vice. Indeed, in its most aggressive incarnations, political perfectionism even countenances “coercive moral paternalism,” “the legal enforcement of morals or the punishment of those who commit

⁴⁶ For more on this distinction, see Thomas Hurka, “Indirect Perfectionism: Kymlicka on Liberal Neutrality,” *Journal of Political Philosophy* 3 (1995): 36-57.

⁴⁷ See, e.g., Ronald Dworkin, *Sovereign Virtue* (Cambridge, Mass.: Harvard University Press, 2000), ch. 6.

⁴⁸ See, e.g., John Finnis, “Is Natural Law Theory Compatible with Limited Government?,” in *Natural Law, Liberalism, and Morality*, ed. Robert P. George (Oxford: Oxford University Press, 1996), 3.

morals offenses.”⁴⁹ To license such regulation, however, is by no means to mandate it; a perfectionist government may always decide to stay its hand and tolerate certain moral transgressions for the sake of other important considerations. But when political perfectionism decides to abide immoral conduct, it does so in deference not to the requirements of liberal public justification, but to the dictates of prudence, the imperative to implement policies that are efficacious as well as moral. Contemporary liberalism, by contrast, imposes a much more peremptory prohibition against governmental moralizing.

V. TOWARD A PERFECTIONIST ACCOUNT OF ANTIDISCRIMINATION LAW

The contrast between contemporary liberalism and (political) perfectionism has significant implications for antidiscrimination law. For whereas perfectionism may invoke any moral principles to condemn discrimination and legitimate antidiscrimination law, liberalism must hew strictly to the terms of public justification recognized by the correct liberal doctrine of political legitimacy. If we cannot condemn discrimination without transgressing these limits, then contemporary liberalism will fail to legitimize at least certain applications of antidiscrimination law. However immoral, discrimination must also violate shared principles of public justification if contemporary liberalism is to legitimate antidiscrimination law.

As with governmental regulation of any wrongful behavior, then, liberals must distinguish two issues when assessing antidiscrimination law: that of the wrongfulness of discrimination based on race or sex and that of the legitimacy of governmental regulation of that wrong. They may not, in other words, uncritically assume that antidiscrimination law effortlessly clears the hurdles liberalism erects in the path of all exercises of political power. Rather, fundamental liberal principles demand greater scrutiny of the justifications

⁴⁹ Robert P. George, *Making Men Moral* (Oxford: Oxford University Press, 1993), 1.

of antidiscrimination law that liberals suppose their doctrines of political legitimacy to entail.

This thesis aims to provide that scrutiny. In particular, Chapters 2, 3, and 4 consider three prominent contemporary liberal doctrines of political legitimacy (respectively, Rawls's "political liberalism," an antiperfectionist version of John Stuart Mill's "harm principle," and Raz's perfectionist version of the harm principle). Each chapter reconstructs the doctrine's case against discrimination based on race or sex and then assesses the robustness of that case, its ability to legitimize various applications of antidiscrimination law. It turns out that each of the doctrines actually thwarts antidiscrimination law in certain circumstances. To be sure, the doctrines underwrite antidiscrimination regimes of varying scopes and strengths, but they all nevertheless fail to justify applications of antidiscrimination law that we typically assume to be perfectly legitimate. This failure, in turn, owes not to the doctrines' assorted idiosyncrasies, but rather to their common contemporary liberal core—namely, their rigid prioritization of personal autonomy over other moral values. Unless we are willing to relax this priority, unless we are willing, in particular, occasionally to infringe personal autonomy for the sake of cultivating moral character, we will have to abide many instances of discrimination.

How should we respond to these shortcomings in the three liberal doctrines of political legitimacy? Some liberals might be willing to bite the bullet, conceding that the doctrines do indeed preclude certain applications of antidiscrimination law but nonetheless refusing to repudiate them. Chapter 5, by contrast, argues that such persistence is misguided. While by no means refuting contemporary liberalism, the chapter attempts to deprive the ideology of much of its allure by developing an alternative account of discrimination based on race or sex that condemns every instance of the

conduct—even those instances condoned by the liberal doctrines of political legitimacy. According to that account, discrimination is wrong not just because it tends to have deleterious consequences for its victims, but also (and primarily) because it necessarily taints the moral characters of its perpetrators. For in forming and acting on an intention to treat another person less favorably because of her race or sex, one fails to accord her the respect to which she is entitled simply in virtue of being a person. Only by espousing this understanding of discrimination can we vindicate the widely shared intuition that the conduct is always wrong.

To condemn discrimination, however, is not to license the government to remedy it. Rather, to legitimize antidiscrimination law, we must first demonstrate that the government's authority reaches the kind of wrong wrought by discrimination. Chapter 6 undertakes this task, constructing a respect-enforcing doctrine of political legitimacy that subjects at least certain kinds of disrespectful conduct to governmental regulation. This doctrine unequivocally forsakes contemporary liberalism for perfectionism, spurning liberal public justification and putting autonomy on a par with other values rather than granting it ascendancy, as most contemporary liberals do. Unlike Raz's liberal perfectionism, moreover, the doctrine is perfectionist not only in its theoretical foundations, but also in its political program—not only in its endorsement of a conception of the good life, but also in its authorization of moralizing governmental regulation. Many liberals will deem this embrace of legal moralism very costly indeed. But it is a cost we must be willing to bear if we are to justify the full extent of modern antidiscrimination regimes.⁵⁰

⁵⁰ The central argument of this thesis differs in at least three fundamental respects from Kimberly Yuracko's attempt to develop a perfectionist account of certain aspects of antidiscrimination law. First, Yuracko considers only discrimination based on sex, ignoring discrimination based on race. Second, she focuses not on the legitimacy of the core provisions of antidiscrimination law, but rather on various court decisions applying the legislation in what she concedes are relatively peripheral cases. Third, and most important, her perfectionism seeks to promote only the flourishing of discrimination's (female) victims

rather than that of its perpetrators. More specifically, her perfectionism seeks not to reform the moral characters of discriminatory employers, but to foster a work environment in which female workers can achieve fulfilment. See Kimberly A. Yuracko, "Private Nurses and Playboy Bunnies: Explaining Permissible Sex Discrimination," *California Law Review* 92 (2004): 147-213; "Sameness, Subordination, and Perfectionism: Toward a More Complete Theory of Employment Discrimination Law," *San Diego Law Review* 43 (2006): 857-97. For a fuller exposition of Yuracko's "feminist perfectionism," see her *Perfectionism and Contemporary Feminist Values* (Bloomington, IN: Indiana University Press, 2003).

Part I

**Contemporary Liberalism and
Antidiscrimination Law**

Chapter 1

Liberal Legitimacy and Public Policy

I. ABSTRACT PRINCIPLES, CONCRETE POLICIES, AND LIBERAL LEGITIMACY

Before proceeding with the preliminary, critical argument of this thesis, we must confront a threshold objection to the polemical strategy that it employs. Recall that this thesis first proposes to assess three of the most prominent contemporary liberal doctrines of political legitimacy for their ability to justify not just the abstract moral principles underlying antidiscrimination law, but also the legislation's concrete applications. It challenges contemporary liberalism, in other words, to vindicate our intuitions about particular applications of antidiscrimination law as well as our intuitions about the general moral status of modern antidiscrimination regimes. According to this challenge, if contemporary liberalism condemns any individual application of antidiscrimination law—especially any application we consider central to the antidiscrimination project—it will have belied its moral promise and forfeited its claim to our allegiance. To prove that its ideal moral principles legitimate actual public policies—that is the challenge this thesis sets for contemporary liberalism.

We might worry that this challenge saddles contemporary liberalism with an undue justificatory burden, expecting a degree of correspondence between the ideology's main tenets and particular governmental actions that no abstract theory can hope to achieve. For in initially formulating its legitimacy criterion, any doctrine of political legitimacy will have to overlook myriad contingencies: facts about political practice, institutional behavior, human nature, and so on. These contingencies, in turn, can prevent the government from strictly complying with any abstract principle. Unless doctrines of political legitimacy may refine their requirements to allow public policies to

accommodate such practical considerations, they might end up precluding so much governmental activity that they disqualify themselves as plausible accounts of legitimate political power.

To avoid this result, we might be inclined to re-construe doctrines of political legitimacy to be more flexible. More specifically, we might consider the possibility that doctrines of political legitimacy can function as two-level principles, articulating their requirements primarily in terms of general goals for the government to promote and only secondarily in terms of specific rules for it to obey. On such an interpretation, a doctrine's first-order rules do not derive directly from its second-order goals. That is, a two-level doctrine does not simply enjoin the government to obey its legitimacy criterion on every occasion. Rather, such a doctrine formulates more nuanced first-order rules that prescribe governmental actions designed to maximize the realization of its second-order goals over the long run, even if some of the prescribed actions violate the express terms of its legitimacy criterion on particular occasions. Only by adopting this indirect approach, the argument goes, can any doctrine of political legitimacy hope to implement its ideal principles in actual circumstances.

If we conceive of doctrines of political legitimacy as two-level principles, then we must reassess this thesis's polemical strategy against contemporary liberalism. That strategy will condemn any doctrine of political legitimacy that forestalls particular applications of antidiscrimination law. But it seems we cannot employ the same strategy and level the same condemnation against a two-level doctrine of political legitimacy. That is, we apparently cannot impugn a two-level doctrine of political legitimacy simply by demonstrating an incongruity between its legitimacy criterion and particular applications of antidiscrimination law—an incongruity that might reflect not some illicit exercise of political power, but a reasonable response to practical constraints. Rather, we apparently

can reject a two-level doctrine of political legitimacy only if it legitimates so many governmental actions that, on the whole, society's main political institutions contravene even the general spirit of its legitimacy criterion.

Not every doctrine of political legitimacy, however, can be construed as a two-level principle. On the contrary, this chapter will argue that contemporary liberal doctrines of political legitimacy defy such an interpretation and thus are indeed vulnerable to this thesis's polemical strategy. For such doctrines are concerned not just with the effects of governmental regulation, but also with the reasons on which it is based. More specifically, the doctrines are antiperfectionist, excluding from the public domain reasons that contravene contemporary liberalism's conception of the person as autonomous. Having denied the legitimacy of reliance on perfectionist considerations, an antiperfectionist doctrine may not prescribe rules for governmental conduct that incorporate those same considerations, lest it license the government to act on the very reasons it was devised to exclude in the first place. And yet, two-level doctrines of political legitimacy are plagued by precisely this incoherence, ostensibly proscribing certain justifications for governmental action only to end up permitting the government to invoke those justifications for the sake of expediency. If contemporary liberal doctrines of political legitimacy take their own restrictions on political power seriously, then they must enforce their requirements consistently. And if contemporary liberal doctrines of political legitimacy cannot function as two-level principles, then they cannot evade this thesis's polemical strategy after all.

II. LIBERAL DOCTRINES OF POLITICAL LEGITIMACY: DEONTIC OR TELIC?

To function as a two-level principle, a doctrine of political legitimacy must assume a particular form. More specifically, all two-level doctrines of political legitimacy

must also be “telic” or “teleological” principles rather than “deontic” or “deontological” ones. According to John Rawls, in teleological moral theories,

the good is defined independently from the right, and then the right is defined as that which maximizes the good. More precisely, those institutions and acts are right which of the available alternatives produce the most good, or at least as much good as any of the other institutions and acts open as real possibilities.¹

To appreciate this way of drawing the telic/deontic distinction, we must consider the theoretically prior distinction between “the good” and “the right.” Whereas a theory of the good is “a view about which properties we ought to want realized in our actions or in the world more generally”—that is, which properties are valuable—a theory of the right is “a view . . . about what individual and institutional agents should do by way of responding to valuable properties.”² A teleological³ theory enjoins agents to respond to valuable properties by seeking to “promote” them unqualifiedly, but to “honour” them “only so far as honouring them is part of promoting them, or is necessary in order to promote them.”⁴ That is, a teleological theory identifies the “right” action in any given situation as the one that will most fully realize the relevant value, even if performing the action requires the agent to transgress the very value the action promises to promote. For example, assuming truthfulness to be valuable, a teleological theory would require a person to lie on a particular occasion if, by so doing, he could induce a greater number of people to be truthful. For a teleological theory, it is the overall condition of a value, not any particular agent’s conformity with it, that defines rules of right action.

Deontological moral theories reject at least one of the tenets of teleology, refusing either to “specify the good independently from the right” or to “interpret the right as

¹ John Rawls, *A Theory of Justice*, rev. ed. (Cambridge, Mass.: Harvard University Press, 1999), 21-22.

² Philip Pettit, “Consequentialism,” in *A Companion to Ethics*, ed. Peter Singer (Oxford: Blackwell, 1991), 230.

³ Pettit considers “consequentialism,” which is one particularly prominent variant of teleology.

⁴ Pettit, “Consequentialism,” 231.

maximizing the good.”⁵ The versions of deontology most relevant for this chapter’s argument are those that reject the second of these tenets, versions that hold that, even if values can in principle be promoted, “at least some values call to be honoured whether or not they are thereby promoted.”⁶ Thus understood, deontological theories posit certain absolute constraints on human action, rules that agents must never violate, even in pursuit of “ends that are morally admirable, or morally obligatory.”⁷ More specifically, deontological theories view certain actions as intrinsically wrong, so that agents must eschew such actions “even when they foresee that their refusal to do such things will clearly result in greater harm (or less good).”⁸ Deontology thus entails results that are anathema to teleology, at least occasionally requiring agents to forgo actions that maximize valuable properties.

We can apply the distinction between teleology and deontology not only to comprehensive theories, but also to discrete principles, such as doctrines of political legitimacy. Thus, whereas a telic principle will countenance only those actions that maximize a particular value, a deontic principle will at least sometimes proscribe such actions, requiring agents to perform alternative ones that realize the value less fully, if at all. Similarly, whereas a telic doctrine of political legitimacy will enjoin the government to “promote” a particular value, a deontic doctrine of political legitimacy will require the government to “honor” that value, even when doing so realizes the value less fully than possible.

Taking contemporary liberalism’s disavowals of teleology⁹ at face value, the Introduction cast at least two of the three kinds of doctrines of political legitimacy in

⁵ Rawls, *Theory of Justice*, 26.

⁶ Pettit, “Consequentialism,” 231.

⁷ Nancy (Ann) Davis, “Contemporary Deontology,” in *A Companion to Ethics*, ed. Peter Singer (Oxford: Blackwell, 1991), 205.

⁸ *Ibid.*, 206.

⁹ See, e.g., Rawls’s critique of utilitarianism in *Theory of Justice*, 24, 26.

deontic, rather than telic, terms. Direct ends-limiting doctrines seem to conform most closely to the deontic mold, authorizing the government to regulate only behavior that produces a certain kind of effect and therefore enjoining it from regulating other behavior that does not—regardless of the desirable consequences it might otherwise achieve. Even though they address themselves in the first instance to society’s main political institutions rather than individual exercises of political power, indirect ends-limiting doctrines can also be understood to impose deontic restrictions on governmental action. Consider, for example, Rawls’s “liberal principle of legitimacy,” according to which “our exercise of political power is fully proper only when it is exercised in accordance with a constitution the essentials of which all citizens as free and equal may reasonably be expected to endorse in the light of principles and ideals acceptable to their common human reason.”¹⁰ This doctrine functions as a deontic principle, imposing categorical, if indirect, constraints on governmental action. For the “exercise of political power” *constitutes* political institutions; it does not merely exhibit their character. That is, by altering the manner in which it exercises political power, the government also alters the very institutions it comprises. To forbid political institutions to assume a particular form is thus to forbid the government to exercise political power in a manner that would cause political institutions to assume that proscribed form. Although it initially imposes categorical constraints on the design of political institutions, Rawls’s liberal principle of legitimacy ends up entailing similarly categorical constraints on the exercise of political power. These latter constraints render the doctrine a deontic principle.

Whereas direct and indirect ends-limiting doctrines of political legitimacy seem to embody the logic of deontology, means-limiting doctrines can more plausibly be understood as telic principles. To be sure, like their ends-limiting counterparts, means-

¹⁰ John Rawls, *Political Liberalism*, paperback ed. (New York: Columbia University Press, 1996), 137.

limiting doctrines impose restrictions on the exercise of political power. But unlike the restrictions imposed by ends-limiting doctrines, the ones imposed by means-limiting doctrines have a *pro tanto* character: they are valid not intrinsically, but only insofar as they actually promote a given end. More specifically, means-limiting doctrines specify some value to be promoted—personal autonomy, in the case of liberal means-limiting doctrines—and then forbid the government to adopt any means in the pursuit of its other ends that would frustrate that value. The restrictions imposed by means-limiting doctrines thus serve a merely instrumental function, (indirectly) guiding the government toward a specified goal rather than constraining its actions for no ulterior purpose. This kind of goal-driven project is one of the hallmarks of teleology.

Despite their strong affinity with deontology, even ends-limiting doctrines of political legitimacy can be construed as *telic* principles. Indeed, ends-limiting doctrines can be *telic* in precisely the same way that means-limiting doctrines are. Consider, first, direct ends-limiting doctrines, which enumerate the ends the government may legitimately pursue. As depicted above, the enumeration of legitimate ends amounts not to a set of goals to be realized, but rather to a list of permissions and corresponding prohibitions to be obeyed. But direct ends-limiting doctrines are amenable to an alternative interpretation: they can indeed be understood to assign the government certain goals to achieve and only incidentally to prohibit it from pursuing other ends. On this understanding, in limiting the government to specific, enumerated ends, direct ends-limiting doctrines do not necessarily condemn the other, prohibited ends as inherently illegitimate—the kind of verdict that a deontic doctrine of political legitimacy would render. Rather, they exclude the other, forbidden ends from the government’s agenda in order to promote the enumerated ones more effectively. Insofar as the prohibitions they

impose thus have merely instrumental, rather than intrinsic, value, direct ends-limiting doctrines can be understood in telic terms.

Indirect ends-limiting doctrines are likewise amenable to a telic interpretation. As exemplified by Rawls's "liberal principle of legitimacy," such doctrines establish criteria to govern the arrangement of society's main political institutions and then evaluate the exercise of political power for its conformity with whatever institutional arrangement the criteria prescribe. As we saw above, if its institutional criteria are cast as categorical constraints on the form political institutions may assume, then an indirect ends-limiting doctrine of political legitimacy is deontic and will countenance the exercise of political power only insofar as political institutions assume an acceptable form. But if its institutional criteria are instead cast as goals for political institutions to fulfill, then such a doctrine is telic and will countenance the exercise of political power insofar as it aids political institutions in realizing their assigned goals. There is, then, no feature inherent in the different kinds of doctrines of political legitimacy that precludes our interpreting them as telic, rather than deontic, principles.

III. TWO-LEVEL DOCTRINES: AN ANALOGY WITH RULE CONSEQUENTIALISM

If we interpret the different kinds of doctrines of political legitimacy as telic principles, then it seems that their teleology is indirect. For rather than directly enjoining the government to promote some abstract value through each of its actions, they prescribe more concrete rules for the government to follow. By following these rules, the doctrines suppose, the government will promote the specified value more fully than it would if it attempted to determine for itself which course of action would be most effective on each particular occasion. Of course, because they prescind from the unique circumstances presented by any given situation, the rules will occasionally misfire, directing the

government to perform an action that turns out to promote the specified value less fully than some alternative action that the government could have identified through its own, direct calculus. But telic doctrines will nevertheless retain the rules and insist that the government follow them so long as compliance with them produces more effective outcomes overall, if not on every particular occasion. Telic doctrines will, in other words, function as two-level principles, whose first-order rules deviate from their overarching values.

In assuming this two-level structure, telic doctrines of political legitimacy resemble indirect forms of consequentialism—most notably, rule consequentialism. Rule consequentialism opposes itself to act consequentialism, the moral theory that holds that “the moral rightness of an act depends entirely on whether the act’s consequences are at least as good as that of any alternative act.”¹¹ Act consequentialism thus derives a rule of conduct (“act in the way that produces the best consequences”) directly from its criterion for assessing states of affairs. By contrast, rule consequentialism contends that “better consequences will result if people do *not* try always to decide what to do by calculating consequences than if they try always to decide in this way.”¹² The theory thus enjoins individual agents not to perform the act that will produce the best consequences on any given occasion, but rather to follow the set of rules that, when also followed by a sufficient number of other people, will produce the best consequences over time. It assesses acts, in other words, not directly in terms of their consequences, but indirectly in terms of their conformity with a particular set of rules, deeming them to be “permissible if and only if [they are] allowed by a code [of rules] that could reasonably be expected to result in as much good as could reasonably be expected to result from any other

¹¹ Brad Hooker, “Rule-Consequentialism,” in *The Blackwell Guide to Ethical Theory*, ed. Hugh LaFollette (Oxford: Blackwell, 2000), 183.

¹² *Ibid.*, 187.

identifiable code.”¹³ It is this two-level structure, this bifurcation of criteria for assessing states of affairs and rules for guiding conduct, that distinguishes rule consequentialism from act consequentialism.

Rule consequentialism’s two levels can differ not only in terms of their practical specificity, but also in terms of their ideological character. Indeed, rule consequentialists insist that their first-order rules of conduct can have the categorical force of deontological rules yet ultimately depend on consequentialist justifications. For example, rule consequentialists might justify the general practice of punishment in consequentialist terms even as they espouse—and demand rigid adherence to—traditional, deontological rules of criminal liability (e.g., the rule prohibiting the punishment of the innocent).¹⁴ So long as better consequences will result if people generally comply with deontological rules than if they perform their own consequentialist calculations, rule consequentialists will espouse deontology as the proper guide for individual conduct.

Just as rule consequentialism’s first-order rules can belie its criteria for assessing states of affairs, so, too, can a telic doctrine of political legitimacy countenance exercises of political power that ostensibly violate its legitimacy criterion. Consider the harm principle, the direct ends-limiting doctrine of political legitimacy that authorizes the government only to prevent harm.¹⁵ Understood in deontological terms, the harm principle prohibits the government from regulating any harmless conduct. But suppose we instead construe the doctrine as a two-level telic principle, one that espouses the goal of minimizing harm but does not enjoin the government to promote that goal directly. So understood, the harm principle could prescribe rules that (at least occasionally) permit the

¹³ Ibid., 188.

¹⁴ See John Rawls, “Two Concepts of Rules,” *Philosophical Review* 16 (1955): 4-13. Cf. Hooker, “Rule Consequentialism,” 188 (“[I]f there will be greater overall good where people are largely disposed to focus and act on non-consequentialist considerations, then consequentialism itself endorses such dispositions.”).

¹⁵ As we shall see in Chapter 4, the harm principle can also function as a means-limiting doctrine of political legitimacy, authorizing the government to use some means—say, coercion—only to prevent harm.

government to regulate harmless conduct—so long as such rules form part of an institutional system that prevents at least as much harm as any alternative system would. Like rule consequentialism, then, a telic doctrine of political legitimacy will tolerate a significant discontinuity between its fundamental principles and its first-order rules.

Among these first-order rules are the rules composing antidiscrimination law. As an exercise of political power, antidiscrimination law must comply with the correct doctrine of political legitimacy's requirements for governmental regulation of private conduct. But if the doctrine assumes a telic, two-level form, then it will require not that each discrete application of antidiscrimination law satisfy its legitimacy criterion, but rather that the legislation, considered as a system of rules, does so. The harm principle, for instance, will countenance applications of antidiscrimination law even to harmless instances of discrimination so long as those applications form part of a regulatory scheme that prevents more harm than any alternative one. If strict governmental compliance with its legitimacy criterion would thwart its broader aims, then a telic doctrine of political legitimacy will have no compunction about relaxing that criterion and authorizing applications of antidiscrimination law that it would otherwise seem expressly to proscribe.

There is at least some reason to suppose that the government will indeed thwart a telic doctrine of political legitimacy's broader aims by strictly complying with its legitimacy criterion. Most notably, like the act-consequentialist injunction to maximize the good, the injunction, say, to prevent harm might be self-defeating. That is, in deliberately attempting to prevent harm on every particular occasion, the government might in fact prevent less harm overall than it otherwise would. Perhaps the government "frequently lack[s] information about the probable consequences of" people's conduct,¹⁶

¹⁶ Hooker, "Rule-Consequentialism," 187-88.

or perhaps it tends to miscalculate the harmfulness of individual conduct even when it has such information.¹⁷ But whatever the exact cause, it is by no means implausible to suppose that the government will prevent harm less effectively if it attempts to identify harmful conduct for itself than if it follows rules designed to maximize harm prevention over time. A telic, two-level version of the harm principle will thus avoid simply instructing the government to regulate only harmful discrimination. Rather, assuming that most instances of discrimination are harmful, it will promulgate rules that authorize the government to regulate all instances of discrimination, harmful and harmless alike. Although such rules nominally contravene the harm principle's legitimacy criterion, they form part of a system that prevents more harm than any practical alternative and thus accord with a telic version of the harm principle.

In the context of antidiscrimination law, a telic doctrine of political legitimacy will countenance departures from its legitimacy criterion for reasons not only of governmental competence, but also of political expediency. Assume that only some instances of discrimination are harmful—in particular, those instances where the victims suffer from some preexisting disadvantage. In such circumstances, it seems that the harm principle would require a bifurcated antidiscrimination regime, one that prohibited discrimination against only disadvantaged racial minorities and women while tolerating discrimination against their more fortunate counterparts. But it is very unlikely that such a regime would ever command the popular support necessary to be enacted (let alone enforced), given the widespread intuition that every instance of discrimination, no matter how harmless, is morally wrong. Thus, even if the government could distinguish between harmful and harmless instances of discrimination, a telic, two-level version of the harm principle would authorize the government to regulate both—a concession to the political

¹⁷ Cf. *ibid.*

realities constraining any doctrine of political legitimacy's implementation of nondiscrimination norms.

By recognizing these kinds of prudential and political considerations, telic doctrines of political legitimacy introduce a significant asymmetry between their legitimacy criteria and the rules of antidiscrimination law. For the latter cannot strictly comply with the former if the doctrines are maximally to realize their fundamental aims. And yet, this thesis's polemical strategy expects a near-perfect symmetry between the two, demanding that the doctrines justify each application of antidiscrimination law not according to extraneous prudential and political considerations, but only according to the explicit terms of their legitimacy criteria. Understood in telic terms, the doctrines seem immune to this strategy. That is, we cannot impugn a telic doctrine of political legitimacy simply by identifying applications of antidiscrimination law that violate its legitimacy criterion, violations that might reflect the doctrine's incorporation of various prudential considerations rather than any deeper incoherence. Instead, we must show that the entire antidiscrimination regime fails maximally to realize its fundamental aim—that, for instance, current antidiscrimination law, which forbids harmful and harmless discrimination alike, prevents less harm than would a more nuanced regime. Far from exposing some deep inconsistency between contemporary liberalism's commitments to limited government and nondiscrimination, it seems that this thesis promises only to confirm the ideology's due regard for prudence.

IV. TWO-LEVEL DOCTRINES AND ANTIPERFECTIONIST LIBERALISM

The preceding discussion has assumed that we can construe contemporary liberal doctrines of political legitimacy as two-level principles. While the previous section demonstrated how we could formally read the different kinds of doctrines in those terms, it did not consider whether more fundamental liberal principles preclude such a reading.

In fact, insofar as contemporary liberal doctrines of political legitimacy espouse antiperfectionism, they cannot assume a two-level form. For antiperfectionism is concerned not just with the consequences of governmental regulation, but also with the reasons for which the government acts.¹⁸ In particular, antiperfectionism prohibits the government from acting for reasons rooted in a particular conception of the good life. Two-level doctrines fail to enforce this prohibition, admitting reasons at the level of first-order rules that are ostensibly forbidden at the level of second-order goals. Put simply, antiperfectionist two-level doctrines are incoherent.

To begin to understand this critique of two-level interpretations of antiperfectionist doctrines of political legitimacy, consider Richard Arneson's critique of rule consequentialism, the paradigmatic two-level moral theory. While recognizing the possibility of multi-tiered theories like rule consequentialism, Arneson nevertheless insists "that the tiers must cohere. Considerations taken to be morally significant in themselves at one level should not be undercut or vanish entirely at other levels."¹⁹ Arneson derives this requirement from a conception of the relationship between moral theory and human agency. In particular, Arneson rejects any radical asymmetry between the rules that govern people's conduct and the fundamental principles that justify those rules—the exact kind of asymmetry that characterizes rule consequentialism. According to Arneson, rule consequentialism "aims to cater to [people's] deontological and common-sense moral judgments about cases without giving these deontological and common-sense elements any place in the statement of fundamental principles."²⁰ The

¹⁸ To be sure, there are some forms of antiperfectionism that *do* care about the consequences of governmental regulation—namely, those forms that require "neutrality of effect." But neutrality of effect does not appear to be a particularly realistic ideal, and most liberals have disavowed it. See, e.g., Rawls, *Political Liberalism*, 192-94.

¹⁹ Richard Arneson, "Sophisticated Rule Consequentialism," *Philosophical Issues* 15 (2005): 245.

²⁰ *Ibid.*, 244.

result, Arneson concludes, is “an unsustainable hybrid strategy,”²¹ a Janus-faced theory that permits people to act on certain kinds of considerations that it derides as misguided.

Put another way, Arneson seems to believe that rule consequentialism holds moral theory hostage to people’s misguided moral beliefs. If consequentialism provides the correct criterion for assessing states of affairs, Arneson wonders, then why should we attempt to tailor first-order rules of conduct to people’s common-sense, deontological judgments about how they should act in particular circumstances? In seeking to have it both ways, to espouse a consequentialist criterion for assessing states of affairs without having to disabuse people of their deontological moral intuitions, rule consequentialism confronts a dilemma:

Either the deontological and other common-sense judgments that we are trying to accommodate should be accepted, in which case they should be reflected directly into the formulation of fundamental principle, or they should be resisted and not accepted, explained away as an understandable mistake to which we are prone. But in that latter case the deontological intuitions pose no bar to acceptance of straight old-fashioned act consequentialism.²²

Rule consequentialism denies this dilemma, claiming a supposed middle ground where abstract consequentialist principles and concrete deontological judgments can coexist. For Arneson, the problem with such a compromise is that it sacrifices moral truth for moral conservatism: rather than attempt to alter the moral status quo, which reflects a prevailing popular commitment to deontology, rule consequentialism entrenches it. It is incoherent, Arneson contends, for a moral theory so to modify its requirements in order to accommodate moral error.

Two-level liberal doctrines of political legitimacy—or at least antiperfectionist ones—suffer from a similar kind of incoherence. Antiperfectionism forbids the government to act for certain kinds of reasons—namely, reasons rooted in a particular conception of the good life. And yet, in an effort to accommodate practical realities and

²¹ Ibid.

²² Ibid., 244-45.

people's moral intuitions, two-level antiperfectionist doctrines of political legitimacy end up permitting the government to act on precisely the perfectionist reasons they purport to proscribe. Although the latter kind of accommodation does so more directly than the former, both kinds violate fundamental antiperfectionist principles.²³

Consider, first, the attempt by an antiperfectionist doctrine of political legitimacy—say, a direct ends-limiting doctrine like the harm principle—to accommodate people's intuition that all instances of discrimination are immoral. If we suppose, once again, that only some instances of discrimination harm their victims, a two-level version of the harm principle will permit the government to proscribe all instances of discrimination, harmful and harmless alike, on the grounds that a more qualified proscription would fail to garner sufficient political support, given people's intuition that discrimination is always immoral. Like rule consequentialism's compromise with people's deontological intuitions regarding right conduct, this compromise between antiperfectionism and people's legal-moralist intuitions regarding discrimination privileges the moral status quo over moral truth (according to antiperfectionism), bending the first-order rules regulating governmental action to fit people's particular conceptions of the good. Such accommodation is incoherent: if antiperfectionism truly provides the correct account of political legitimacy, then it should insist on the validity of that account rather than tailor public policy according to the kinds of considerations it was supposed to exclude in the first place. Indeed, attempts by two-level antiperfectionist doctrines of political legitimacy to accommodate people's moralistic intuitions are not merely incoherent, but *unjust*, for antiperfectionism purports to enforce some of the most important political protections to which persons are entitled in virtue of their status as autonomous beings. To relax these protections by permitting the government to act on

²³ The following paragraphs preview some of the issues considered in greater depth in Chapter 3.

considerations rooted in people's conceptions of the good is to violate contemporary liberalism's foundational commitments.

Even attempts by two-level antiperfectionist doctrines of political legitimacy to accommodate more practical considerations can lead to incoherence and injustice. Consider once again a two-level version of the harm principle, which would authorize the government to proscribe harmful and harmless instances of discrimination alike if it would prevent less harm overall by attempting to distinguish between harmful and harmless instances of discrimination than it would simply by proscribing the conduct wholesale. The problem with this indirect approach to preventing harm is that it privileges the efficient realization of a teleological goal—harm prevention—over antiperfectionism's essentially deontological limits on the exercise of political power. An antiperfectionist version of the harm principle limits the government to preventing harm because the concept of harm ostensibly transcends competing conceptions of the good: all conceptions agree that the government should prevent harm, whatever else it should do, and all largely agree in their determinations of what conduct is actually harmful.²⁴ By hewing to the terms of this agreement, the harm principle treats all conceptions of the good and their adherents fairly, in accordance with the liberal conception of persons as autonomous. The harm principle risks subverting this agreement, and therefore treating conceptions of the good and their adherents unjustly, by extending itself to encompass harmless conduct, even for the sake of preventing harmful conduct more efficiently. For even if conceptions of the good generally agree about what conduct is harmful, they will disagree about just what costs to other values the government may legitimately incur in order to prevent harm. The government will have to perform this calculation as it decides whether to proscribe harmless conduct in the name of preventing more harm overall,

²⁴ This agreement resembles Rawls's notion of an "overlapping consensus" of conceptions of the good. See Rawls, *Political Liberalism*, 133-72.

inevitably contradicting the contrary assessments of at least some conceptions of the good in the process. Therefore, a two-level version of the harm principle will force the government to take a position on certain questions of the good life. Because antiperfectionism forbids the government to take such a position, a two-level version of a direct ends-limiting doctrine of political legitimacy like the harm principle is incoherent. And because antiperfectionism purports to embody fundamental considerations of justice, such a doctrine is unjust.

This is not to suggest that antiperfectionism precludes every attempt to expand a doctrine of political legitimacy to encompass forms of behavior that the doctrine nominally prohibits the government from regulating—that, say, literally every instance of a given form of conduct must be harmful before the harm principle will authorize the government to prohibit the conduct wholesale. For example, an antiperfectionist version of the harm principle can coherently authorize the government to enact wholesale proscriptions of a given form of conduct based on empirical generalizations about the conduct's tendency to cause harm, even if some of the conduct's instantiations prove harmless. Unlike a strategy of accommodating people's intuitions about the wrongfulness of harmless conduct or a strategy of treating harmful and harmless instances of a given form of conduct alike in order to avoid the difficulty of distinguishing between the two, a strategy of relying on generalizations about the likely consequences of a given form of conduct need not require the government to take a position on questions concerning the good life. To be sure, the specter of perfectionism will haunt even this latter strategy, since conceptions of the good will inevitably disagree about just how likely a given form of conduct must be to cause harm before the government may prohibit all its instantiations. Nevertheless, just as there will be a broad range of agreement among conceptions of the good about what forms of conduct are always harmful, so, too, will

there be a broad range of agreement among conceptions of the good about what forms of conduct are sufficiently likely to cause harm that the government may prohibit all their instantiations, harmful and harmless alike. So long as a doctrine of political legitimacy permits the government to legislate only within this broad range of agreement, it can avoid any governmental entanglement with a particular conception of the good and thus the incoherence that characterizes the kinds of two-level strategies considered above.

It might be objected that the preceding critique of two-level versions of antiperfectionist doctrines of political legitimacy wrongly assumes that one- and two-level versions respond to the same problem of political morality, when, in fact, they respond to different problems and should therefore be assessed according to different criteria. We can better appreciate this objection by once again considering the analogy with rule consequentialism. To those theorists who, like Arneson, argue that rule consequentialism is incoherent, Tyler Cowen has responded that act consequentialism and rule consequentialism are mutually consistent responses to different problems in moral theory rather than genuine competitors and, consequently, that the latter does not face the dilemma of either collapsing into the former or descending into incoherence.²⁵ According to Cowen, although both variants of consequentialism seek to maximize the same specified end (say, utility, in the case of utilitarianism), each does so in response to a different “maximization problem” with a different set of “constraints.” In particular, where act consequentialism considers the problem of a single agent deciding how to act on a single occasion, rule consequentialism considers the problem of “bundling” individual actions “across time,” of “picking some policy today, under the constraint that a similar policy (or rule) be applied in subsequent time periods.”²⁶ Act consequentialism and rule consequentialism, in other words, employ the same maximizing reasoning to

²⁵ Tyler Cowen, “Rule Consequentialism Makes Sense After All,” *Social Philosophy and Policy* 28 (2011): 226.

²⁶ *Ibid.*

guide human action; rule consequentialism simply labors under an additional decisional constraint from which act consequentialism is free. On this view, rule consequentialism is not some incoherent compromise between consequentialism and deontology, but rather a reasonable attempt to apply consequentialist reasoning under a precisely defined set of conditions.

Cowen's defense of rule consequentialism suggests an analogous defense of two-level versions of antiperfectionist doctrines of political legitimacy. According to this defense, single-level versions of antiperfectionist doctrines of political legitimacy instruct the government how to act only on single occasions, making no further assumptions about what implications these isolated actions might have for the government's subsequent conduct. Two-level versions, by contrast, instruct the government how to act on the assumption that it must act more or less consistently over time—that, if it adopts a particular policy on one occasion, then it must continue to follow that policy in similar circumstances. Both one- and two-level versions thus seek to enforce the logic of antiperfectionism, preventing the government from intentionally favoring a particular conception of the good life; two-level versions simply do so over the course of the government's actions rather than on just one particular occasion.

This argument appears to absolve two-level versions of antiperfectionist doctrines of political legitimacy of the charge of incoherence. Recall that, according to that charge, two-level versions are incoherent because they end up permitting the government to act on the very reasons that they were designed to exclude in the first place. It now seems that the situation is more complex. In comprehending the full course of governmental action, rather than just discrete governmental acts, two-level versions are forced to adopt a diachronic perspective, considering the implications of discrete governmental acts for the government's future compliance with the antiperfectionist mandate. And in adopting this

perspective, two-level versions provide the government with a reason that their synchronic, one-level counterparts do not—namely, a reason to ensure its compliance with antiperfectionism over time, not just on any particular occasion. Sometimes, the government will best honor this reason by acting on perfectionist considerations. Thus, even if a two-level version of an antiperfectionist doctrine of political legitimacy sometimes authorizes the government to act for perfectionist reasons, the government will still be acting for the *further* reason of complying with antiperfectionism. The government, in other words, will have a reason to act for perfectionist reasons only insofar as it has a further reason to comply with antiperfectionism.

The problem with this argument is that it does not fully appreciate the significance of the government's acting on a reason rooted in a particular conception of the good life. Suppose that, because of people's existing attitudes, the government could best comply with antiperfectionism over the long run by banning extramarital sexual conduct today—a prohibition that almost all antiperfectionists would consider anathema. Although the government, in enacting this prohibition, would be acting on perfectionist reasons (*viz.*, people's beliefs about the immorality of extramarital sexual relations), it would be doing so only for the sake of the further reason of complying with antiperfectionism. And yet, an antiperfectionist would still have reason to object to the prohibition. For the government does not merely foresee and accept that the prohibition will favor one particular conception of the good life, but rather deliberately adopts the conception as part of its justification for the policy. It thereby treats the conception as legitimate grounds for regulation, as within the proper bounds of public reason. To be sure, the conception is made relevant to the justification for the prohibition only by a further set of reasons that is external to the conception and thus remains thoroughly antiperfectionist. But the existence of a further set of antiperfectionist reasons cannot overcome the fact that the

government has lent its imprimatur to a particular conception of the good, an endorsement that contravenes antiperfectionism.

We have seen how direct ends-limiting doctrines of political legitimacy, such as the harm principle, cannot coherently assume a two-level form. A similar incoherence plagues two-level versions of indirect ends-limiting doctrines. To be sure, in the first instance, such doctrines prescribe rules not for particular exercises of political power, but rather for the general structure of society's main institutions. We thus might be tempted to think that they are more amenable than direct ends-limiting doctrines to a two-level interpretation. More specifically, because they adopt a systemic perspective that appears to focus less on the reasons for which the government acts on any given occasion than on the outcomes produced by the general structure of institutions over time, it might seem that indirect ends-limiting doctrines can more readily accommodate first-order rules that countenance periodic departures from antiperfectionism.

It turns out, however, that two-level indirect ends-limiting doctrines prove every bit as incoherent as direct ones. For despite their systemic focus, such doctrines seek to arrange institutions to instantiate the very same principle espoused by direct ends-limiting doctrines: the antiperfectionist principle forbidding the government to act for reasons rooted in a particular conception of the good life.²⁷ Ignoring this principle in the assessment of individual exercises of political power while upholding it in the design of institutions is just as incoherent as ignoring it in the assessment of individual exercises of political power while upholding it in the specification of the government's legitimate ends.

What about means-limiting doctrines of political legitimacy? Surely they can assume a two-level form? After all, they are not, strictly speaking, antiperfectionist,

²⁷ Cf. Thomas Nagel, "Justice and Nature," *Oxford Journal of Legal Studies* 17 (1997): 303-04 (arguing that, despite its systemic focus, a liberal theory of justice can remain deontological rather than consequentialist).

leaving the government free to act on whatever valid reasons it chooses so long as it employs only permissible means in doing so. Means-limiting doctrines thus seem immune to the charge of incoherence when they relax their first-order rules better to realize their fundamental aims. It is difficult, however, to draw such a general conclusion about means-limiting doctrines, which vary widely in terms of the liberalism of their foundational principles and thus in terms of the stringency of their restrictions on political power. It seems that the more a means-limiting doctrine's foundational principles resemble antiperfectionist ones—that is, the more the doctrine privileges autonomy over other values and precludes moralistic policies—the less able it will be coherently to adopt a two-level structure. Indeed, a means-limiting doctrine that grounds its limits on political power exclusively in the value of autonomy will likely prove so stringent that it effectively functions as a direct ends-limiting doctrine, nominally permitting the government to advert to perfectionist considerations but effectively rendering such considerations otiose. And a means-limiting doctrine that has this effect will suffer from the same incoherence that characterizes two-level ends-limiting doctrines if it promulgates first-order rules that depart too significantly from its fundamental principles.

V. ANTIPERFECTIONISM AND THE RULE OF LAW

In casting doubt on antiperfectionism's ability to accommodate two-level doctrines of political legitimacy, this chapter has thus far focused on only two sets of normative considerations: (1) fundamental principles of political morality (here, antiperfectionism) and (2) the more specific moral rules that directly regulate the government's exercise of political power. It has argued that the latter must accord with the spirit of the former, that a doctrine of political legitimacy cannot coherently espouse antiperfectionism in one breath and countenance perfectionist policies in the next one. And yet, it might be objected that this argument ignores a third set of normative

considerations that might require antidiscrimination law to diverge from the form we would expect it to assume were it attempting strictly to comply with antiperfectionism. More specifically, antidiscrimination law is composed not of moral rules, but of *legal* rules. We thus might think it unreasonable to demand the same correspondence between fundamental principles and the rules of antidiscrimination law that we expect between fundamental principles and first-order moral rules. For unlike moral rules, legal rules must satisfy the criteria of the “rule of law.” Foremost among these criteria is generality, the requirement that laws be cast in terms capacious enough to encompass a relatively broad range of cases rather than just some narrow subset.²⁸ This requirement promotes one of the most important benefits of the rule of law: providing individuals with adequate guidance so that they can conform their conduct to the law’s dictates. If laws were insufficiently general, making liability turn on myriad contingencies peculiar to particular cases, individuals would not be able to determine in advance whether they might incur liability by performing a particular action. Consider a law that forbids not consumer fraud simpliciter, but only fraud against consumers who happen to be particularly vulnerable, in the sense that the fraud affects their well-being significantly rather than just marginally. By predicating liability on what will usually be an unpredictable contingency (viz., the magnitude of the fraud’s impact on the victim’s overall well-being), such a law fails to apprise would-be fraudsters of the legal risks they incur by performing a particular action. Similarly, we might think that antidiscrimination law cannot, consistent with the generality requirement, proscribe only those instances of discriminatory treatment that harm their victims to a certain degree, lest discriminators incur liability for mere lack of prescience.

²⁸ See Lon L. Fuller, *The Morality of Law*, rev. ed. (New Haven, Conn.: Yale University Press, 1969), 46-49.

If the rule-of-law requirement of generality would indeed preclude a prohibition against only harmful discrimination, as opposed to discrimination *tout court*, then it seems that an antiperfectionist doctrine of political legitimacy can coherently justify the full extent of antidiscrimination law after all. So long as at least some instances of discrimination harm their victims, there exists an antiperfectionist rationale for at least a limited prohibition against discrimination. The rule-of-law requirement of generality then demands that we extend this limited prohibition to all discriminatory conduct, harmful and harmless alike, so that would-be discriminators can more readily predict the legal consequences of their conduct. Unlike the two-level strategy considered earlier in this chapter, this one is not incoherent. For whereas the relationship between fundamental principles of political morality and first-order moral rules is *internal*, in the sense that both are grounded in the same kinds of considerations and thus must cohere in spirit, the rule of law imposes *external* constraints on a political morality's pursuit of its legal program. These constraints, moreover, are not perfectionist, but rather are grounded in neutral concepts like individual liberty. Thus, while antiperfectionism cannot coherently evade this thesis's polemical strategy by adopting two-level doctrines of political legitimacy, it apparently can do so by coupling a single-level doctrine with a commitment to the rule of law.

Although there is much truth in this line of argument, it provides less succor to antiperfectionism than we might initially suppose. First, it is far from obvious that a law prohibiting only harmful instances of discrimination would violate the rule of law's generality requirement. Many laws, after all, predicate legal liability on rather unpredictable contingencies. Consider statutory-rape laws, which, at least in many jurisdictions, allow no defense for a mistake concerning the victim's age. Or consider the requirement that a homicide defendant's fear of imminent death or serious bodily injury

have been objectively reasonable in order for him to claim self-defense, where the reasonableness of his belief is assessed by a jury after the fact. In both cases, the law acknowledges that its subjects will often fail to predict all the circumstances that determine their liability, but it imposes liability nevertheless, expecting them to err on the side of caution by eschewing conduct that skirts the legal limits. To be sure, it might turn out to be more difficult to anticipate others' vulnerability to discrimination than, say, the age of one's sexual partner or the lethality of one's adversary; socioeconomic disadvantage might be less apparent than youth or malicious intent. But antidiscrimination law can mitigate this unpredictability short of proscribing every instance of discrimination. It can, for instance, require a discrimination plaintiff to prove that the alleged discriminator had reason to know that his conduct would cause harm (because, say, the plaintiff's application for employment bespoke a disadvantaged background). This procedural device reduces the chances that a defendant will be found liable based on contingencies he had no ability to foresee—one of the primary aims of the rule of law. It is thus too quick to conclude that the rule of law's generality criterion permits—indeed, requires—antiperfectionism to enact a broader prohibition against discrimination than it could of its own accord. Before taking such a drastic step, we must consider other, more modest measures that provide adequate guidance to would-be discriminators without circumventing antiperfectionism's strictures.

Second, even if a prohibition against only harmful discrimination would violate the rule of law, why should rule-of-law considerations necessarily trump a strict adherence to antiperfectionist principles? This objection takes two forms, depending on whether we understand rule-of-law considerations to be internal or external to a given political morality. On the one hand, the rule of law might constitute part of a broader political morality, rather than an independent constraint on it. Rawls, for instance,

enshrines the requirements of the rule of law in his first principle of justice, which, in turn, he derives from his more general liberal commitments.²⁹ In his theory, then, rule-of-law considerations are on a par with other liberal tenets, including antiperfectionism. Rawls provides no clear decision rule for resolving conflicts among these tenets, and given their apparent incommensurability, it is difficult to see how such a rule could even be devised. Nevertheless, if antiperfectionism allows antidiscrimination law to prohibit only harmful instances of discrimination, while the rule of law requires it to prohibit every instance, contemporary liberalism should provide some account of why it chooses to privilege the latter liberal tenet over the former one.

On the other hand, the rule of law might impose external constraints on antiperfectionism; law might have its own intrinsic morality, independent of any particular political theory.³⁰ Although this possibility would obviate the need for a decision rule to adjudicate conflicts between antiperfectionism and the rule of law (like any other political morality that sought to use the law to achieve its goals, antiperfectionism would have to play by the law's own rules), it would not allay the concern that antiperfectionism was relying on the rule of law to break free from its own constraints. So long as antiperfectionism could adduce some justification for even a few core applications of antidiscrimination law, it could then invoke the independent requirements of the rule of law to bootstrap those applications into a comprehensive antidiscrimination regime, even one identical to what perfectionism might entail. Antiperfectionism should not be able to avoid its unpalatable implications so easily. To be sure, we must allow antiperfectionism some degree of flexibility in implementing its requirements. Thus, when most instances of discrimination tend to be harmful, it is only reasonable to allow antiperfectionism to rely on the rule of law to reach the other,

²⁹ See Rawls, *Political Liberalism*, 291. See also Rawls, *Theory of Justice*, 206-07.

³⁰ See generally Fuller, *Morality of Law*.

harmless ones. But when empirical conditions change so that most instances of discrimination are harmless, such reliance seems strained, a case of the proverbial tail wagging the dog. To determine whether the rule of law merely subserves antiperfectionism or instead threatens to efface it, we must first assess the independent strength of antiperfectionist justifications for antidiscrimination law in different empirical conditions. Only when the justifications encompass a sufficient number of instances of discrimination on their own strength can the rule of law then help antiperfectionism to support the kind of comprehensive antidiscrimination regimes contemporary liberal democracies now employ.

VI. CONCLUSION

We might think that liberal doctrines of political legitimacy can evade this thesis's polemical strategy by adopting a two-level structure, espousing unequivocally antiperfectionist fundamental principles but prescribing first-order rules that at least occasionally countenance perfectionist policies. This chapter has attempted to foreclose this proposed evasion. While liberal doctrines of political legitimacy can nominally be read in telic terms, they cannot assume a two-level form, lest they end up descending into incoherence by permitting the government to act on the very reasons they were designed to exclude. Because they take antiperfectionism's reason-restricting function as their *raison d'être*, ends-limiting doctrines—whether direct or indirect—will always resist a two-level interpretation. Means-limiting doctrines are somewhat more flexible, but insofar as they espouse strongly liberal foundational principles, their two-level iterations also risk incoherence. With its focus on the reasons underlying exercises of political power, antiperfectionism severely limits the kinds of considerations to which the government may advert and thus the extent to which the rules regulating governmental conduct may diverge from more fundamental liberal tenets.

For antiperfectionism to preclude two-level principles is for it to demand a relatively close correspondence between its foundational principles and actual public policies. If antiperfectionism were instead to tolerate public policies that incorporated considerations it ostensibly forbade, then it would provide no genuine alternative to perfectionism. It thus seems perfectly reasonable to expect antiperfectionism to justify those public policies, like antidiscrimination law, whose legitimacy we would never think to call into question—the very expectation underlying this thesis’s polemical strategy.

The expectation of congruence between actual policies and fundamental principles seems especially appropriate in the case of antidiscrimination law, a policy that we view as embodying some of our most confident moral beliefs. To be sure, we distinguish between morality and its institutionalization, recognizing that public policy must account for various practical factors if it is to be effective and that it will thus never perfectly instantiate morality. But at least in the case of antidiscrimination law, we take public policy as a baseline against which we evaluate competing moral theories. It is to this standard that the next three chapters will hold three of the most prominent contemporary liberal doctrines of political legitimacy.

Chapter 2

Political Liberalism

I. SYSTEMIC DISCRIMINATION AND INDIVIDUAL LIBERTY

We tend to vacillate between two alternative paradigms for comprehending and responding to discrimination based on race or sex. Whereas we sometimes attribute discrimination to the unjust structure of society's main political, social, and economic institutions, we often focus instead on the discrete discriminatory actions of society's individual members. In construing discrimination, then, we opt either for a systemic account or for an individualistic one, for an account that situates individual discriminatory acts in a broader institutional context or for one that largely abstracts from that context, considering each act alone. Much depends on which account we espouse. For when we conceive of discrimination as the upshot of an unjust institutional structure, we tend to direct our remedial efforts—at least in the first instance—at the structure itself, not at the individual persons who operate within its scope. When, by contrast, we trace discrimination directly to the unjust actions of individual persons, those persons, not institutions, tend to bear the brunt of our reformatory zeal. The two accounts of discrimination presuppose two dichotomous conceptions of responsibility for injustice and therefore entail two dichotomous ameliorative programs.

Of course, though perhaps in tension with each other, the two accounts of discrimination need not be mutually exclusive. In particular, we can explicate discrimination systemically without having to forgo all moral evaluation of individual discriminatory acts; though it focuses on institutions, the systemic account does not necessarily exculpate people for their wrongdoing. Where the systemic account of discrimination differs from the individualistic one is in viewing institutional

discrimination as a whole greater than the sum of its parts, as a unique injustice that does not simply reduce to individual discriminatory acts. While not necessarily downplaying individual responsibility for injustice, such an account nevertheless displaces our moral appraisal and censure from persons to institutions.

Despite the compatibility of the two accounts of discrimination, most contemporary liberal political philosophers emphasize the systemic account over the individualistic one, considering discrete discriminatory acts largely for their broader, institutional consequences. Thomas Nagel, for example, attributes the political and economic disadvantage suffered by black Americans not to individual instances of discrimination, but to the functioning of “the great American caste system.”¹ Indeed, he goes so far as to assimilate such disadvantage with socioeconomic inequality more generally, inequality he considers to be “a pervasive and fundamental feature of the system.”² To be sure, Nagel recognizes that any instance of discrimination perpetrates “an *additional* injustice,” beyond any material disadvantage, against its victim.³ But he nonetheless trains his fire on society’s main institutions and advocates public policies—like race-based affirmative action—designed to counteract those institutions’ pernicious effects.

John Rawls assesses discrimination from a similarly systemic perspective, likening race and sex to other human “contingencies” that social institutions tend to translate into socioeconomic inequality.⁴ Only by targeting this synergism between personal characteristics and institutional structure, Rawls suggests, can we begin to make our society more just. On Rawls’s view, as on Nagel’s, we should combat discrimination

¹ Thomas Nagel, “John Rawls and Affirmative Action,” *Journal of Blacks in Higher Education* 39 (2003): 84.

² Thomas Nagel, “Equal Treatment and Compensatory Discrimination,” *Philosophy & Public Affairs* 2 (1973): 353.

³ *Ibid.*, 360. See also Nagel, “Affirmative Action,” 84.

⁴ John Rawls, *Justice as Fairness*, ed. Erin Kelly (Cambridge, Mass.: Harvard University Press, 2001), 64-65.

not just by thwarting and punishing individual discriminators, but primarily by reconfiguring the institutional arrangements that abet and compound their wrongdoing.

As this brief examination of Rawls's and Nagel's approaches to discrimination makes clear, not only do the systemic accounts of discrimination adopted by contemporary liberal political philosophers focus on the broader, institutional effects of the conduct; they focus on the effects discrimination has for one value in particular: socioeconomic equality. On this view, discrimination is unjust insofar as it disadvantages entire groups of citizens vis-à-vis their compatriots, and antidiscrimination law is one important means of preserving socioeconomic equality. Systemic in form, contemporary liberal accounts of discrimination are also egalitarian in substance.

Contemporary liberal political theories face a significant challenge in balancing this systemic egalitarianism with their antiperfectionism. As we saw in the Introduction, contemporary liberalism conceives of the person as an autonomous being. Liberals suppose this conception, in turn, to require that the government heed the strictures of public reason, justifying its policies not according to any particular conception of the good life, but rather according to principles everyone can reasonably accept.⁵ To ensure governmental compliance with this requirement, a liberal political theory will have to enact a system of rights and liberties that insulates certain personal conduct from governmental interference, conduct that is constitutive of autonomy and whose regulation

⁵ Like the rendition of antiperfectionism in the Introduction and Chapter 1, this one focuses on the *reasons* for which the government acts rather than the *consequences* of governmental action. It thus might not seem to fit Rawls's political theory. Joseph Raz, for one, has ascribed to Rawls a version of "comprehensive neutrality," according to which "[o]ne of the main goals of governmental authority, which is lexically prior to any other, is to ensure for all persons an equal ability to pursue in their lives and promote in their societies any ideal of the good of their choosing" (Joseph Raz, *The Morality of Freedom* (Oxford: Oxford University Press, 1986), 115. See *ibid.*, 118 ("If Rawls' theory is based on a doctrine of neutrality it is a doctrine of comprehensive neutrality.")). According to this understanding of antiperfectionism, the government may not take actions that have the effect of furthering a particular conception of the good. However accurately Raz's understanding of antiperfectionism might capture earlier iterations of Rawls's theory, it does not survive the transition from Rawls's original theory of justice to political liberalism. In his later work, Rawls explicitly disavows "neutrality of effect or influence" and instead espouses "neutrality of aim," which holds "that the state is not to do anything intended to favor or promote any particular comprehensive doctrine rather than another, or to give greater assistance to those who pursue it" (John Rawls, *Political Liberalism*, paperback ed. (New York: Columbia University Press, 1996), 193).

thus cannot, in most instances, be justified according to public reason. But such conduct will also tend to have significant systemic effects, and tolerating it will tend to impede the realization of systemic values, including socioeconomic equality. The challenge for contemporary liberalism, then, is to explain how it will resolve the inevitable conflicts between its antiperfectionist commitment to individual autonomy and its systemic account of socioeconomic equality.

This challenge is especially acute for Rawls's theory of "political liberalism," a systemic account of justice that Rawls portrays as both liberal and egalitarian. It is liberal, Rawls claims, because it seeks to safeguard individual autonomy by guaranteeing basic liberties, and it is egalitarian because it requires the eradication of much socioeconomic inequality.⁶ But as Nagel notes, Rawls does not obfuscate the relationship between the two values of liberty and equality. On the contrary, seeking to preserve "each individual's sovereignty over himself," Rawls breaks decisively with "movements farther to the left" and proclaims his theory unequivocally liberal by establishing a "strict priority of individual rights and liberties over the reduction of social and economic inequalities," a ranking of moral objectives that Nagel considers to be "the true core of liberalism."⁷ Because the priority of liberty still permits the government to remedy at least "excessive social economic inequalities," it does not, Rawls insists, render his theory "libertarian."⁸ Nevertheless, it does often hamstring the government's efforts to curb inequality and thus reflects Rawls's thoroughgoing liberalism.

⁶ Rawls, *Political Liberalism*, 6-7. For Rawls's enumeration of the other egalitarian features of his theory, see *ibid.* Although these features might not qualify Rawls's theory as "strictly egalitarian" (see Raz, *Morality of Freedom*, 225-27), a conception of equality nevertheless plays a foundational role in that theory. For more on the relationship between Rawls's theory and egalitarianism, see Samuel Scheffler, "What Is Egalitarianism?," *Philosophy & Public Affairs* 31 (2003): 5-39; and Norman Daniels, "Democratic Equality: Rawls's Complex Egalitarianism," in *The Cambridge Companion to Rawls*, ed. Samuel Freeman (Cambridge: Cambridge University Press, 2003).

⁷ Thomas Nagel, "Rawls and Liberalism," in *The Cambridge Companion to Rawls*, ed. Samuel Freeman (Cambridge: Cambridge University Press, 2003), 64-65, 66-67.

⁸ Rawls, *Political Liberalism*, lviii.

The priority of liberty casts serious doubt on the legitimacy of central facets of antidiscrimination law. As we have seen, Rawls adopts a systemic account of discrimination, condemning the conduct for its inegalitarian effects. But Rawls also subordinates his egalitarianism to his antiperfectionist commitment to individual autonomy, a hierarchy codified in the priority of liberty. Rawls's theory will legitimate antidiscrimination law, then, only to the extent that the legislation does not infringe the rights and liberties assigned priority over socioeconomic equality. That is, his theory will legitimate antidiscrimination law only to the extent that his antiperfectionism does not thwart his egalitarianism.

This chapter argues that the conception of justice expounded by Rawls in *Political Liberalism* privileges its antiperfectionism over its egalitarianism to such an extent that it condemns much antidiscrimination law as illegitimate. More specifically, this chapter contends that the priority of liberty—particularly as applied to freedom of association—forestalls many core applications of antidiscrimination law if we view discrimination based on race or sex as a systemic, inegalitarian injustice, as Rawls's theory evidently does. Taken together, political liberalism's antiperfectionist commitment to individual autonomy and systemic account of socioeconomic equality belie the structure of much contemporary antidiscrimination law, an incongruity that exposes some of the theory's most significant shortcomings.

II. THE AIMS AND LIMITS OF POLITICAL LIBERALISM

Political Liberalism and Public Policy

Before proceeding with this argument, we must assure ourselves that this chapter has not set for Rawls's theory a challenge it need not meet. Chapter 1 considered—and ultimately rejected—the suggestion that a liberal doctrine of political legitimacy could accommodate any gap between its fundamental principles and actual public policies

simply by adopting a two-level structure. Antiperfectionism, it was argued, precludes such a structure, lest its doctrines of political legitimacy end up permitting the government to act on the very reasons they were designed to exclude in the first place. That argument, however, assumed that antiperfectionist liberalism actually purports to justify existing public policies, that it seeks to account not just for the abstract principles underlying contemporary western liberal democracies, but also for those societies' particular laws and institutions. If a theory disclaimed this ambition, it seems it would be impertinent to fault the theory for any gap between its fundamental principles and actual public policies.

Rawls appears to limit the ambitions of his theory in precisely this way, explicitly circumscribing its subject matter and scope of application. For example, he defends his “emphasis on the Reformation and the long controversy about toleration as the origin of liberalism” and his exclusion of the seemingly more urgent contemporary problems of racism and sexism, noting that his theory seeks “to offer an account of political and social justice that is more satisfactory than the leading familiar traditional conceptions” and thus “limit[s] itself . . . to a family of classical problems that had been at the center of the historical debates concerning the moral and political structure of the modern democratic state.”⁹ Moreover, regardless of its precise subject matter, political liberalism purports to govern not every particular policy enactment, but rather the general arrangement of society's main institutions.¹⁰ While political liberalism should thus specify the minimum, baseline criteria any social institution must satisfy, it seems it need indicate neither the actual institutions society must establish nor the precise policies those institutions must legislate.

⁹ Ibid., xxx.

¹⁰ Ibid., 257-58.

And yet, to accept this limited conception of Rawls's agenda would be to leave untapped much of his theory's prescriptive potential—potential Rawls himself recognizes and even touts. Despite initially restricting his subject matter to “a few long-standing classical problems” in political philosophy, Rawls proceeds on the “assumption . . . that a conception of justice worked up by focusing on [those problems] should be correct, or at least provide guidelines for addressing further questions,” fully expecting that “once we get the conceptions and principles right for the basic historical questions, those conceptions and principles should be widely applicable to our own problems also.”¹¹ Indeed, Rawls supposes that his theory of justice is “complete,” that its main principles, together with “guidelines of public inquiry,” will yield answers “to all, or to nearly all,” of the most pressing questions of “basic justice”¹²—questions that surely must include that of discrimination based on race or sex. Given this assumption, Rawls must also believe that his theory's principles provide significant guidance not just for the design of abstract institutions, but also for the formulation of concrete policies; hence his elaboration of a “four-stage sequence” of different points of view from which the members of a given society are to consider various questions of justice, with each successive point of view admitting more empirical information about the society than did the previous, more abstract one.¹³ It is not unreasonable, then, for us to expect Rawls's principles of justice, duly supplemented by relevant empirical facts, to apply to particular questions of public policy—especially one as urgent as that of discrimination based on race or sex.¹⁴

¹¹ Ibid., xxxi.

¹² Ibid., 225.

¹³ See John Rawls, *A Theory of Justice*, rev. ed. (Cambridge, Mass.: Harvard University Press, 1999), 171-76. See also Rawls, *Political Liberalism*, 397-99.

¹⁴ Rawls himself illustrates how we might apply his principles of justice at the “legislative stage” of his “four-stage sequence” to the issue of health care. See Rawls, *Justice as Fairness*, 173-75.

To be sure, we cannot expect political liberalism to settle every policy question. For not every policy question implicates justice, and even those policy questions that do must often be resolved not through the direct application of principles of justice, but through political compromise, a process that inevitably alloys justice with other, less lofty considerations. If we are to assess Rawls's theory of justice on its own terms, then, we should be wary of following philosophers like Thomas Pogge who attempt to "develop[] [the theory] in a way that renders it more concrete and makes its full progressive potential more visible,"¹⁵ lest in so doing we unwittingly distort it beyond its original purposes. But we can heed this warning without having to prescind in our philosophical reflections from all policy questions, some of which bear the hallmarks of the principles of justice on their face. Discrimination epitomizes this latter type of policy question and thus serves as a particularly important test case for Rawls's theory of justice.

Discrimination and Ideal Theory

We thus should hesitate neither to apply Rawls's theory of justice to at least some concrete questions of public policy nor to expect it to provide significant guidance for resolving them. But is the issue of discrimination in particular the kind of issue that is amenable to a Rawlsian analysis? Once again, in demarcating the limits of his theory, Rawls seems to suggest not. In particular, Rawls conceives of his project as primarily one of "ideal" or "strict compliance" theory, which "assumes strict compliance [with the demands of justice] and works out the principles that characterize a well-ordered society under favorable circumstances."¹⁶ "Non-ideal" or "partial compliance" theory, by contrast, "studies the principles that govern how we are to deal with injustice."¹⁷ Whereas nonideal theory grapples with the many profound injustices currently afflicting our societies, ideal

¹⁵ Thomas W. Pogge, *Realizing Rawls* (Ithaca: Cornell University Press, 1989), 4.

¹⁶ Rawls, *Theory of Justice*, 216.

¹⁷ *Ibid.*, 8.

theory largely disregards them, envisioning circumstances in which they all have been eradicated.

Given the distinction between ideal and nonideal theory, it seems that the latter fully comprehends the issue of discrimination, obviating almost any role for Rawls's ideal principles of justice in an account of antidiscrimination law. For it seems we must assume that the citizens of Rawls's ideal society simply do not discriminate based on race or sex; as Amy Gutmann avers, "[i]t would be a blatant contradiction for a political philosophy to posit an ideal society that is beset by a legacy of racial injustice."¹⁸ Therefore, if we construe Rawls's principles of justice as formulated solely for an ideal society, we must refrain from applying them directly to the injustice of discrimination. Rather, we must develop alternative principles that, while perhaps informed by, are nonetheless distinct from Rawls's ideal ones, principles that are specially tailored to societies that comply only partially with the demands of justice.¹⁹ Indeed, Rawls himself explicitly consigns "[t]he serious problems arising from existing discrimination and distinctions based on gender and race" to nonideal theory, strongly implying that such problems necessitate a unique set of principles not derived directly from his ideal ones.²⁰ In proposing to assess the ability of Rawls's ideal principles of justice to justify antidiscrimination law, we seem not to be taking Rawls's theory on its own terms.

Whatever the precise boundaries of ideal theory, however, we should expect Rawls's theory of justice to address discrimination. For, at least in his later work, Rawls adulterates his ideal "well-ordered society" with an ample dose of realism, admitting to its

¹⁸ Amy Gutmann, "Responding to Racial Injustice," in K. Anthony Appiah and Amy Gutmann, *Color Conscious: The Political Morality of Race* (Princeton: Princeton University Press, 1996), 109. On one reading, Gutmann's argument might be question-begging, since Rawls's principles of justice apply not to citizens directly, but only to the "basic structure" of society. This aspect of Rawls's theory has elicited trenchant criticism. See, e.g., G.A. Cohen, "Where the Action Is: On the Site of Distributive Justice," *Philosophy & Public Affairs* 26 (1997): 3-30.

¹⁹ Cf. Gutmann, "Responding," 109.

²⁰ Rawls, *Justice as Fairness*, 66.

ranks people who, despite their other virtues, are patently unjust. As Rawls explains, while the main institutions of a well-ordered society are “effectively regulated” by a “reasonable liberal conception of justice,” every citizen need not endorse it. Rather, so long as the citizens who do endorse the conception “are in an enduring majority,” the society qualifies as well-ordered and just, remaining within the purview of ideal theory.²¹ We can reasonably envisage Rawls’s well-ordered society, then, as one whose citizenry includes some recalcitrant racists and sexists but whose other, tolerant citizens are nonetheless politically ascendant. And, in such circumstances, we can reasonably insist that Rawls’s ideal principles of justice legitimate antidiscrimination law.²²

Even in the context of societies that do not qualify as well-ordered, Rawls’s ideal principles of justice are not otiose. For example, Rawls maintains that the priority of liberty over questions of socioeconomic equality holds not only in well-ordered, fully just societies, but also in any society that enjoys “reasonably favorable conditions,” any society characterized by “social circumstances which, provided the political will exists, permit the effective establishment and the full exercise” of the basic liberties.²³ To emphasize that such a society need not be fully just, Rawls cites the contemporary United States as an example—despite its many flagrant violations of his principles of justice.²⁴ In seeking to bring unjust societies in closer conformity with the principles of justice, Rawls suggests, we must observe at least some of the limits established by the principles themselves; the principles of justice govern not only the well-ordered society, but also, at

²¹ Rawls, *Political Liberalism*, xlix-l.

²² Indeed, even in his earlier work, Rawls acknowledged that unjust conduct would persist in conditions of strict compliance. See, e.g., Rawls, *Theory of Justice*, 212 (“[W]e need an account of penal sanctions however limited even for ideal theory. Given the normal conditions of human life, some such arrangements are necessary.”). See also Robert S. Taylor, “Hate Speech, the Priority of Liberty, and the Temptations of Nonideal Theory,” *Ethical Theory and Moral Practice* (forthcoming 2011): 6. (“Strict compliance must be consistent with some level of illegal behavior (like illicit discrimination); otherwise, ideal theory would be utopian, as some crime is an unavoidable feature of human societies.”).

²³ Rawls, *Political Liberalism*, 297.

²⁴ *Ibid.*

least to some degree, the path leading to it.²⁵ We thus can classify the issue of discrimination squarely within nonideal theory without having to forgo all appeals to the principles of justice when assessing antidiscrimination law. The challenge for Rawls's theory of justice to justify antidiscrimination law still stands.

III. THE LIBERAL PRINCIPLE OF LEGITIMACY

As we saw in the Introduction, contemporary liberal political philosophers take as their primary task the specification of the conditions under which the exercise of political power is legitimate, a task that Rawls, too, places at the center of his theory. Again, as the Introduction explained, many liberals perform this task by establishing criteria that a society's main institutions must satisfy to wield political power legitimately. And likewise, Rawls joins these other liberals in adopting such an institutional focus. He thus considers "the first subject of justice" to be society's "basic structure," "the way in which the major social institutions fit together into one system, and how they assign fundamental rights and duties and shape the division of advantages that arises through social cooperation."²⁶ Not the actions of individual persons or the internal proceedings of private associations, but the structural arrangement of society's main institutions—this is the central focus of Rawls's political liberalism.

In developing a systemic theory of justice rather than an individualistic one, Rawls hopes to effect "an institutional division of labor between the basic structure and the rules applying directly to individuals and associations and to be followed by them in particular transactions."²⁷ But in adopting this bifurcated approach, does Rawls not risk

²⁵ Cf. Pogge, *Realizing Rawls*, 127, 136 (considering the relevance of the priority of liberty in nonideal theory).

²⁶ Rawls, *Political Liberalism*, 257-58.

²⁷ *Ibid.*, 268-69. Some theorists criticize this aspect of Rawls's theory. See, e.g., Cohen, "Where the Action Is"; Liam B. Murphy, "Institutions and the Demands of Justice," *Philosophy & Public Affairs* 27 (1998): 251-91. This chapter cannot assess such criticisms, but instead adopts Rawls's own understanding of his theory of justice as one that applies primarily to major social institutions rather than to private associations or individual actions.

subverting his commitment to socioeconomic equality, effectively permitting people to act outside the domain of society's main institutions without regard for how they might disadvantage others?²⁸ Rawls seems willing to assume this risk for the sake of other values—in particular, the values of political legitimacy and individual liberty. As Nagel explains, by accepting the “gulf between the egalitarianism that determines the design of [society's] institutions and the individualism that motivates its members when they act in the context of those institutions,” Rawls evinces “his unqualified attachment to the liberal tradition despite his strong institutional egalitarianism,” an allegiance that entails a commitment to a strict separation of “political theory” and justice, on the one hand, from “personal morality,” on the other.²⁹ The systemic perspective of Rawls's theory of justice thus dovetails with his goal of extricating politics from morality.

Rawls embeds this systemic perspective in the very foundation of his theory of justice, positing as the “fundamental organizing idea” of political liberalism “that of society as a fair system of social cooperation between free and equal persons viewed as fully cooperating members of society over a complete life.”³⁰ As with the other central concepts of political liberalism, this idea is not some metaphysical truth, but rather one of a number of “certain fundamental ideas seen as implicit in the public political culture of a democratic society.”³¹ Citizens in modern democracies reject any conception of their political relations with one another as aristocratic or otherwise hierarchical, conceiving of them instead as subject to “fair terms” and “an idea of reciprocity,” the notion that

²⁸ For one expression of this concern, see Cohen, “Where the Action Is,” 5-10; G.A. Cohen, “Incentives, Inequality, and Community,” in *The Tanner Lectures on Human Values*, Vol. 13, ed. Grethe Peterson (Salt Lake City: University of Utah Press, 1992).

²⁹ Nagel, “Rawls and Liberalism,” 82.

³⁰ Rawls, *Political Liberalism*, 9.

³¹ *Ibid.*, 13.

benefits and burdens should be distributed fairly among all the participants in a cooperative scheme.³²

The idea of society as a fair system of social cooperation informs all the subsequent ideas Rawls employs to elaborate political liberalism, imbuing the entire theory with its institutional perspective. Thus, in seeking a conception of the person commensurate with that idea, Rawls emphasizes the person's role as "a citizen, that is, a normal and fully cooperating member of society over a complete life."³³ According to Rawls, we suppose full participants in the institutions of social cooperation to be free in that they possess "two moral powers (a capacity for a sense of justice and for a conception of the good) and the powers of reason (of judgment, thought, and inference connected with these powers)." Free citizens are also equal insofar as they possess these powers "to the requisite minimum degree to be fully cooperating members of society," entitling them to enjoy equal civic standing among their compatriots.³⁴ In the political realm, one's freedom and equality derive not from some fundamental, extra-institutional capacity, but solely from one's ability to function within the institutions of social cooperation.

To honor the conception of citizens as free and equal participants in social cooperation, the institutions of the basic structure must eschew exercises of political power that violate citizens' freedom or equality. The threat of such illegitimate exercises of political power is particularly acute in modern democratic societies, which are characterized by certain "facts," foremost among which is the "fact of reasonable pluralism," "the diversity of reasonable comprehensive religious, philosophical, and moral doctrines."³⁵ Comprehensive doctrines are life plans that attempt to assimilate

³² Ibid., 15, 16.

³³ Ibid., 18.

³⁴ Ibid., 19.

³⁵ Ibid., 36.

“conceptions of what is of value in human life, and ideals of personal character, as well as ideals of friendship and of familial and associational relationships,” into relatively coherent sets of practical principles that can guide people as they choose among various life options.³⁶ While free citizens, who have the capacity to form conceptions of the good, espouse comprehensive doctrines, they do not all endorse the same ones. On the contrary, modern democratic societies will harbor a wide range of conflicting comprehensive doctrines. This diversity stems largely from the inevitable fallibility of human reasoning, the sources of which Rawls dubs the “burdens of judgment,” “the many hazards involved in the correct (and conscientious) exercise of our powers of reason and judgment in the ordinary course of political life.”³⁷ Free citizens thus disagree in their comprehensive doctrines not necessarily because of bad faith or intellectual shortcomings, but often because of the ineradicable limits of human reason. This latter kind of disagreement is reasonable and “inevitable”—at least in a society that permits freedom of thought and conscience.³⁸

Whereas those citizens who espouse reasonable comprehensive doctrines recognize the “burdens of judgment” and accept the “fact of reasonable pluralism,” adherents to unreasonable ones attribute the plurality of comprehensive doctrines to the inferior intellects or the malignant motives of their ideological opponents.³⁹ They deny that that plurality is reasonable—let alone an ineliminable feature of society—insisting instead that there exists only one comprehensive doctrine “to be recognized by all citizens who are fully reasonable and rational.”⁴⁰ In so doing, Rawls concludes, they ignore important, necessary features of free human reason.

³⁶ Ibid., 13.

³⁷ Ibid., 56. For a list of the burdens of judgment, see *ibid.*, 56-57.

³⁸ Ibid., 37.

³⁹ Ibid., 58.

⁴⁰ Ibid., 134.

Moreover, those who espouse unreasonable comprehensive doctrines will not hesitate to employ the government's power to impose their beliefs on their ideological opponents.⁴¹ Rawls deems such exercises of political power illegitimate, since, according to the "fact of oppression," "a continuing shared understanding on one comprehensive . . . doctrine can be maintained only by the oppressive use of state power."⁴² Were we to engage in such oppression, we would corrupt the fundamental political relationship among citizens in a democratic society, a relationship characterized by two primary features. First, as we have seen, it occurs within the "basic structure" of society—the major political, social, and economic institutions and the ways in which those institutions distribute the benefits and burdens of social cooperation. In addition, because Rawls assumes that we enter the basic structure only by birth and exit it only by death, he considers our subjection to it involuntary.⁴³ The second feature of the relationship among citizens in a democratic society is the unique character of democratically exercised political power. While Rawls defines all political power as "coercive power backed by the government's use of sanctions," he notes that the political power of a democratic society "is ultimately the power of the public, that is, the power of free and equal citizens as a collective body."⁴⁴ By denying the "fact of reasonable pluralism" and seeking to impose their beliefs on their fellow citizens, adherents to unreasonable comprehensive doctrines propose to enlist the institutions of the basic structure and the collective power of their society not for "public" purposes, but for their own private ones. Were unreasonable citizens to succeed in this usurpation, they would corrode the institutional context of social cooperation, violating the freedom and equality of their fellow citizens and rendering any exercise of political power illegitimate.

⁴¹ Ibid., 138.

⁴² Ibid., 37.

⁴³ Ibid., 135-36.

⁴⁴ Ibid., 136.

To distinguish this injustice from the legitimate exercise of political power, Rawls promulgates a doctrine of political legitimacy that he dubs “the liberal principle of legitimacy.” According to that doctrine, “our exercise of political power is fully proper only when it is exercised in accordance with a constitution the essentials of which all citizens as free and equal may reasonably be expected to endorse in the light of principles and ideals acceptable to their common human reason.”⁴⁵ These shared “principles and ideals,” in turn, derive from a “political conception of justice,” a conception that governs only the basic structure of society (as opposed to private associations and the personal choices of individual persons) and that is “freestanding” in that it presupposes no particular comprehensive doctrine.⁴⁶ To be sure, Rawls assumes that an “overlapping consensus” of reasonable comprehensive doctrines will support his preferred political conception of justice, but we must nevertheless be able to defend such a conception without invoking any particular comprehensive doctrine.⁴⁷ According to the liberal principle of legitimacy, only if a constitution is founded on this kind of political conception of justice will it legitimize the exercise of political power.

In espousing this indirect ends-limiting doctrine of political legitimacy, political liberalism confirms both its institutional perspective and its antiperfectionist commitment to public reason. Evincing the former feature, the liberal principle of legitimacy predicates the legitimate exercise of political power on the just arrangement of society’s main *institutions*, a systemic requirement that follows ineluctably from a theory of justice that conceives of society as a “fair system of cooperation” and that takes the “basic structure” as its “first subject.” And by restricting the kinds of reasons we may adduce to justify a particular institutional structure, political liberalism confirms its antiperfectionism, its scrupulous “neutrality of aim,” according to which “basic

⁴⁵ Ibid., 137.

⁴⁶ Ibid., 11, 12.

⁴⁷ Ibid., 15.

institutions and public policy are not to be designed to favor any particular comprehensive doctrine.”⁴⁸ The liberal principle of legitimacy enforces the dictates of public reason, requiring us to resolve fundamental questions of public policy not by advertent to our respective comprehensive doctrines, but rather by “conduct[ing] [our] public political discussions of constitutional essentials and matters of basic justice within the framework of what each [of us] sincerely regards as a reasonable political conception of justice.”⁴⁹

This requirement complicates the case for particular public policies like antidiscrimination law. For if we are to honor the strictures of public reason, we may not adduce just any justification for antidiscrimination law, which is obviously one of the “constitutional essentials” that concerns “matters of basic justice.” In particular, the liberal principle of legitimacy enjoins us from invoking our respective comprehensive doctrines to justify the exercises of political power entailed by the legislation, requiring us instead to formulate our justification in the terms of public reason, which, in turn, derive from the principles of the political conception of justice. Only after examining these principles can we assess the compatibility of antidiscrimination law with political liberalism.

IV. A FALSE START: DISCRIMINATION AND THE “UNREASONABLE”

We might think that we can make quicker work of justifying antidiscrimination law according to Rawls’s political conception of justice. For, we might wonder, do most discriminators not espouse *unreasonable* comprehensive doctrines? In addition to acknowledging the fact of reasonable pluralism and its implications for the exercise of political power, those who espouse reasonable comprehensive doctrines are by definition willing to propose “fair terms of cooperation” to their compatriots, thereby tacitly

⁴⁸ Ibid., 193-94.

⁴⁹ Ibid., 1.

endorsing certain fundamental moral ideas implicit in the public culture of modern democracies, foremost among which is the conception of citizens as free and equal.⁵⁰ Because many discriminators deny that some of their fellow citizens are equal (and possibly even free) members of their political community, they espouse unreasonable comprehensive doctrines. The case for antidiscrimination law now seems to follow ineluctably: antidiscrimination law suppresses certain unreasonable comprehensive doctrines, enforcing the fundamental normative commitments of modern constitutional democracies against those citizens who reject them.

This argument, however, proceeds far too quickly. First, we might challenge the assumption that all discriminators necessarily espouse “unreasonable” comprehensive doctrines, as Rawls understands that term. While, as a matter of psychological and sociological fact, most discriminators may indeed deny the equality of some of their fellow citizens, some may actually discriminate based on race or sex not because they repudiate the main tenets of political liberalism, but rather because they harbor more subtly biased beliefs. For example, it is at least possible for a discriminator to reject the general moral equality of some of his fellow citizens while nonetheless viewing them as free and equal *for political purposes*. He might, in other words, proclaim the inherent superiority of his race or sex without thereby proposing to exclude his putative inferiors from equal participation in the political life of his society or to violate their basic rights and liberties.⁵¹ Perhaps somewhat more realistically, one might discriminate not from some inveterate, ardent hatred, but rather from a more mild (though still morally objectionable) aversion to difference in general. In harboring either of these two more temperate forms of prejudice, a discriminator would render his comprehensive doctrine only *partially* unreasonable, since he would still view all his fellow citizens as free and

⁵⁰ Ibid., 49-50.

⁵¹ Cf. Erin Kelly and Lionel McPherson, “On Tolerating the Unreasonable,” *Journal of Political Philosophy* 9 (2001): 47.

equal at least in the political realm—a possibility Rawls himself recognizes when he notes that “[a] comprehensive doctrine can be unreasonable on one or several issues without being simply unreasonable.”⁵² Unless discriminators’ comprehensive doctrines are fully unreasonable, the government may not seek to suppress them—through antidiscrimination law or by any other means.

Even if we allow that all discriminators necessarily espouse fully unreasonable comprehensive doctrines, denying the political equality of some of their fellow citizens, it still does not follow that we may employ whatever means necessary to combat their odious beliefs. For just like their reasonable counterparts, adherents to unreasonable comprehensive doctrines enjoy all the basic rights and liberties afforded by Rawls’s political conception of justice. As Rawls himself explains, “[t]here is not one account of toleration for reasonable doctrines and another for unreasonable ones. Both cases are settled by the appropriate political principles of justice and the conduct those principles permit.”⁵³ Adherents to unreasonable comprehensive doctrines, in other words, do not forfeit basic rights like freedom of conscience and association simply by espousing odious beliefs, since “[t]he basis of such rights . . . is independent of the content of one’s views.”⁵⁴ In confronting unreasonable citizens, then, we must always respect their rights and liberties, observing the limits on political power imposed by political liberalism.

Of course, these limits are by no means absolute and may be breached in certain circumstances. In particular, Rawls thinks it likely that a modern, pluralistic society will harbor “unreasonable and irrational, and even mad, comprehensive doctrines,” insisting that we must “contain them so that they do not undermine the unity and justice of

⁵² Rawls, *Political Liberalism*, lvi n.31. Cf. Kelly and McPherson, “Tolerating,” 42.

⁵³ John Rawls, *The Law of Peoples* (Cambridge, Mass.: Harvard University Press, 1999), 178-79. See also Rawls, *Political Liberalism*, 64-65; Jonathan Quong, “The Rights of Unreasonable Citizens,” *Journal of Political Philosophy* 12 (2004): 316-17.

⁵⁴ Joshua Cohen, “A More Democratic Liberalism,” *Michigan Law Review* 92 (1994): 1538.

society.”⁵⁵ But in what actions, exactly, does “containing” unreasonable comprehensive doctrines consist? As Jonathan Quong rightly notes, “containment” cannot simply reduce to preventing unreasonable citizens from infringing their compatriots’ rights and liberties, a protective mission that is independently authorized by Rawls’s principles of justice. Quong thus defines “containment” as “any policy whose primary intention is to undermine or restrict the spread of ideas that reject the fundamental tenets of liberal democracy.”⁵⁶ But even this definition leaves many questions unanswered.⁵⁷ How should we understand the notions of “undermining” and “restricting” unreasonable comprehensive doctrines? Surely, we must encroach on the liberty of those who espouse such doctrines to quarantine their ideas, but just how intrusively may we interfere? May we curb unreasonable citizens’ private activities, or may we prevent those citizens only from garnering political power and enshrining their views in the law? And what if an unreasonable comprehensive doctrine is so marginal that it cannot reasonably be said to “undermine the unity and justice of society”? Must the government then afford its adherents significant leeway?

Though Rawls does not answer these questions, he nevertheless suggests that his political conception of justice countenances only more limited interference with the liberty of unreasonable citizens. Recall that Rawls condemns unreasonable comprehensive doctrines primarily for their denial of the fact of reasonable pluralism and their concomitant willingness “to use the public’s political power—a power in which citizens have an equal share—to enforce a view bearing on constitutional essentials about which citizens as reasonable persons are bound to differ uncompromisingly.”⁵⁸ It thus

⁵⁵ Rawls, *Political Liberalism*, xviii-xix. See also *ibid.*, 64 n.19 (comparing the task of “containing” unreasonable doctrines to that of curbing “war and disease”).

⁵⁶ Quong, “Unreasonable Citizens,” 323.

⁵⁷ Quong proceeds to answer many of these questions, but it is unclear whether his conclusions accord with Rawls’s theory.

⁵⁸ Rawls, *Political Liberalism*, 138.

seems that the government may directly target unreasonable comprehensive doctrines only to prevent their adherents from commandeering the state apparatus for their own purposes. Indeed, Rawls even maintains that a society will be “well-ordered”—it will be effectively governed by a political conception of justice—so long as “unreasonable doctrines (these, we assume, always exist) do not gain enough currency to undermine society’s essential justice.”⁵⁹ If an unreasonable doctrine remains relatively marginal in society and lacks the popular support necessary to attain political power, then political liberalism will most likely condemn governmental interference with the liberty of its adherents. To be sure, the government may undoubtedly prevent the doctrine’s adherents from violating the rights of their compatriots. But in such cases, it is not the unreasonableness of the violators’ comprehensive doctrine, but the rights violations themselves that warrant the use of political power.⁶⁰

We thus cannot justify antidiscrimination law according to Rawls’s political conception of justice simply by condemning discriminators’ comprehensive doctrines as unreasonable. Only if the adherents to these doctrines appear likely to enlist the government’s political power to impose their beliefs on their fellow citizens may the government legitimately abridge their liberty. This is not to suggest that the government may not implement regulation to prevent them from committing particular discriminatory acts, regulation whose legitimacy will depend on the proper balance to be struck among political liberalism’s competing normative commitments. Rather, it is only to insist that the government may not suppress a comprehensive doctrine and violate its adherents’ liberty simply because it is unreasonable.

⁵⁹ *Ibid.*, 39.

⁶⁰ This reading seems to accord with Rawls’s discussion of the “toleration of the intolerant” in *A Theory of Justice*, where he is less concerned with questions of political legitimacy and, generally speaking, seems more willing to countenance governmental interference with individual liberty to promote justice. But he nevertheless licenses regulation of “intolerant” groups “only when the tolerant sincerely and with reason believe that their own security and that of the institutions of liberty are in danger” (193).

V. THE INJUSTICE OF DISCRIMINATION

To determine whether Rawls's political conception of justice can accommodate antidiscrimination law, we must examine its account of the wrongfulness of discrimination based on race or sex; simply citing the unreasonableness of discriminators' comprehensive doctrines will not suffice. Political liberalism, however, may not invoke just any reasons to provide this account. In particular, it must advert not to reasons that derive from our comprehensive doctrines, but only to "public reasons" rooted in the political conception of justice. For, as we have seen, the liberal principle of legitimacy requires that the use of political power be justified according to reasons all citizens can reasonably accept, regardless of the particular comprehensive doctrines they happen to espouse. Only if discrimination based on race or sex violates the political conception of justice will political liberalism legitimate antidiscrimination law.

Rawls's political conception of justice posits two principles, each of which discrimination based on race or sex might violate. The first principle guarantees citizens' freedom, holding that "[e]ach person has an equal claim to a fully adequate scheme of equal basic rights and liberties, which scheme is compatible with the same scheme for all; and in this scheme the equal political liberties, and only those liberties, are to be guaranteed their fair value." The second principle requires that socioeconomic inequalities "satisfy two conditions: first, they are to be attached to positions and offices open to all under conditions of fair equality of opportunity; and second, they are to be to the greatest benefit of the least advantaged members of society."⁶¹ If political liberalism, with its liberal principle of legitimacy, is to authorize antidiscrimination law, discrimination based on race or sex must offend at least one of the two principles of justice.

⁶¹ Rawls, *Political Liberalism*, 5-6.

Recall, however, that Rawls's political conception of justice is systemic in that it takes the "basic structure" as its "first subject." Rawls thus presents the two principles of justice not as rules governing individual conduct, but as a response to the question of how "the basic institutions of a constitutional democracy should be arranged if they are to satisfy the fair terms of cooperation between citizens regarded as free and equal."⁶² By so construing the principles, Rawls introduces at least a superficial discontinuity between political liberalism and much contemporary antidiscrimination law, which usually addresses itself expressly to private persons rather than public institutions.⁶³ To ground antidiscrimination law firmly in the principles of justice, then, we must demonstrate how particular instances of discrimination, perpetrated by individual persons, conspire to corrupt the broader institutional environment within which they occur.

Some instances of discrimination have such conspicuous systemic effects that they obviously contravene one or both of the principles of justice. In particular, the principles seem directly to condemn most, if not all, discriminatory acts perpetrated by the government itself, acts that necessarily alter society's main institutions. For example, by instituting slavery, apartheid, a caste system, or de jure segregation, the government would establish major social institutions that deny some of its citizens equal basic rights and liberties, a patent violation of the first principle of justice. The first principle also requires governmental institutions—especially legal institutions—to satisfy the various criteria of the rule of law,⁶⁴ which, Rawls contends, all derive from the ideal of "formal justice" or "justice as regularity"—the principle that laws should be administered impartially and consistently.⁶⁵ Obviously, the government betrays this ideal, and thus violates the first principle of justice, whenever it illegitimately discriminates among its

⁶² Ibid., 4.

⁶³ See, e.g., the U.K. Equality Act 2010, § 13.

⁶⁴ Rawls, *Political Liberalism*, 291.

⁶⁵ Rawls, *Theory of Justice*, 51, 206-07.

citizens in performing its official functions.⁶⁶ Political liberalism thus unequivocally authorizes those facets of antidiscrimination law that proscribe governmental discrimination.

However, because of the relatively limited purview of its principles of justice, political liberalism must develop a more oblique case against discrimination perpetrated by nongovernmental, private institutions like economic firms and private associations. Indeed, Rawls repeatedly emphasizes that the principles of justice directly regulate only the major social and political institutions composing the basic structure and explicitly includes among those institutions the overall “structure of the economy” rather than individual economic firms, consigning the latter to the sphere of the nonpolitical.⁶⁷ If it is to legitimate antidiscrimination law, political liberalism must establish an indirect connection between the discriminatory actions of individual economic firms and the principles of justice governing society’s overall economy.

Any appearance of a more direct connection between discrimination by individual economic firms and the first principle of justice⁶⁸ is illusory—even if we include such firms within the basic structure alongside political institutions. For by discriminating based on race or sex, businesses do not infringe the basic rights and liberties enshrined in the first principle of justice, which include “freedom of thought and liberty of conscience; the political liberties and freedom of association, as well as the freedoms specified by the liberty and integrity of the person; and finally, the rights and liberties covered by the rule of law.”⁶⁹ When, for instance, a restaurant refuses to serve a black person, it deprives him

⁶⁶ See Tommie Shelby, “Race and Social Justice: Rawlsian Considerations,” *Fordham Law Review* 72 (2004): 1705-06.

⁶⁷ Rawls, *Justice as Fairness*, 10. This statement threatens to embroil us in various exegetical controversies. Cohen, for example, claims that Rawls equivocates about which institutions actually constitute the basic structure (“Where the Action Is,” 18-24).

⁶⁸ Nagel, for one, simply asserts that the first principle “obviously” condemns discrimination (“Affirmative Action,” 82).

⁶⁹ Rawls, *Political Liberalism*, 291.

not of any of these rights and liberties, but rather of other kinds of social benefits, benefits that are not guaranteed by the first principle of justice. We thus cannot reasonably construe the first principle as directly authorizing antidiscrimination law in contexts like those of employment and the provision of public accommodations.

Nor can we link discrimination in such contexts more obliquely with the first principle of justice. To be sure, by discriminating based on race or sex, employers and purveyors of public accommodations certainly can foster an institutional environment that erodes the “worth” or “value” of the basic rights and liberties for their victims, the general usefulness of the rights and liberties to the victims in pursuing their various projects. The first principle, however, guarantees the “fair value,” or (relatively) equal worth, not of all the basic rights and liberties, but only of the political liberties, ensuring that people enjoy merely formally equal access to the nonpolitical liberties.⁷⁰ Only if it is relatively pervasive in society, reinforcing other exclusionary social practices, will private discrimination diminish the “fair value” of the political liberties. And because, as we have seen, such discrimination infringes none of the other, nonpolitical rights and liberties, it preserves at least the formal equality of its victims’ access to them.⁷¹ The first principle of justice, then, neither directly nor indirectly prohibits economic firms from discriminating based on race or sex.

The second principle of justice, by contrast, seems to provide a more promising basis for condemning discrimination. For we might reasonably include antidiscrimination law within the “set of rules that govern the transactions and agreements between individuals and associations” and that are enforced by the institutions of the basic structure to preserve “background justice,” the general societal conditions that prevent individual transactions and agreements from upsetting the distributional pattern mandated

⁷⁰ Ibid., 325-27.

⁷¹ See Seana Valentine Shiffrin, “Race, Labor, and the Fair Equality of Opportunity Principle,” *Fordham Law Review* 72 (2004): 1647.

by the second principle of justice.⁷² That principle requires, inter alia, that social “positions and offices”—including private sector jobs—be “open to all under conditions of fair equality of opportunity.” The requirements of fair equality of opportunity, in turn, encompass not just those of merely formal equality of opportunity, which establishes a social system where “careers are open to talents,” where “positions are open to those able and willing to strive for them” so that “all have at least the same legal rights of access to all advantaged social positions.”⁷³ Fair equality of opportunity guarantees everyone not only this formal access to social positions, but also “a fair chance to attain them.” Rawls construes this requirement as a means of ensuring that “those with similar abilities and skills . . . have similar life chances,” that “[t]he expectations of those with the same abilities and aspirations [are not] affected by their social class.”⁷⁴ Therefore, if fair equality of opportunity obtains in a society, someone born to the poorest family should enjoy the same chance of attaining a particular social position as someone born to the wealthiest—so long as they both possess and cultivate the same talents and ambitions.

Thus defined, fair equality of opportunity appears to condemn most instances of discrimination based on race or sex. Indeed, we might even be tempted to conceive of the principle “as an extension of the ideal of nondiscrimination” itself,⁷⁵ a temptation that is especially strong in the contexts of education and employment. Insofar as seats at universities constitute “advantaged social positions,” discrimination in university admissions violates formal equality of opportunity by foreclosing such positions to some people who are perfectly “able and willing to strive for them.” When relatively pervasive,

⁷² See Rawls, *Political Liberalism*, 268.

⁷³ Rawls, *Theory of Justice*, 57, 62.

⁷⁴ *Ibid.*, 63.

⁷⁵ Richard J. Arneson, “Against Rawlsian Equality of Opportunity,” *Philosophical Studies* 93 (1999): 77. Arneson proceeds, however, to reject the principle of fair equality of opportunity as both too narrow and too broad, as both overlooking certain wrongful instances of discrimination and condemning certain other, unobjectionable ones. To maintain its focus on Rawls’s broader account of political legitimacy, this chapter ignores Arneson’s criticisms and assumes, *arguendo*, that the principle of fair equality of opportunity provides a satisfactory account of the wrongfulness of discrimination.

moreover, such discrimination also violates fair equality of opportunity, since it promotes an economic system that hinders some people from gaining the skills and experience necessary for securing jobs and other further advantaged social positions.⁷⁶ Similarly, because discrimination in hiring also excludes its victims from advantaged social positions and, in so doing, diminishes their prospects for attaining further such positions, it violates both formal and fair equality of opportunity. In both contexts, discrimination compromises the institutional preconditions for fair equality of opportunity mandated by Rawls's second principle of justice.

Discrimination, moreover, will violate at least fair equality of opportunity whether it is "direct" or "indirect." According to U.K. antidiscrimination law, one directly discriminates against another person based on race or sex if, "because of" her race or sex, one treats her "less favourably than [one] treats or would treat others."⁷⁷ Such treatment tends to violate both formal and fair equality of opportunity for the reasons just noted. Again according to U.K. antidiscrimination law, one indirectly discriminates against another person based on race or sex if one "applies" to her a "provision, criterion or practice" that one would apply equally to people of other races or the other sex but that nevertheless disproportionately disadvantages members of her race or sex and for which one can offer no compelling race- or sex-neutral justification.⁷⁸ Unlike perpetrators of direct discrimination, perpetrators of indirect discrimination seem to violate only fair equality of opportunity, since the positions to which they attach the disproportionately burdensome requirements or conditions remain at least formally open to members of all races and both sexes. Nevertheless, at least in the context of employment, fair equality of opportunity seems to condemn most instances of direct and indirect discrimination.

⁷⁶ Rawls, *Justice as Fairness*, 11.

⁷⁷ Equality Act 2010, § 13(1).

⁷⁸ *Ibid.*, § 19.

It is less obvious, however, how fair equality of opportunity condemns discrimination in the provision of public accommodations. In particular, it seems that we must contort the meaning of the term “advantaged social position” beyond recognition to encompass the benefits provided by restaurants, hotels, and other public accommodations. Therefore, if we are to condemn discrimination in the provision of public accommodations as a violation of fair equality of opportunity, we must demonstrate a more indirect connection between such discrimination and the purpose fair equality of opportunity ultimately purports to serve—namely, the maintenance of “background justice.” Such a connection seems relatively easy to draw. For a citizen will most likely be unable to function as a fully cooperating member of society and to pursue “advantaged social positions” unhindered if he is systematically excluded from the public accommodations whose benefits his compatriots freely enjoy, a form of systemic disadvantage that at least indirectly violates fair equality of opportunity. Fair equality of opportunity thus seems to condemn discrimination in all the contemporary contexts governed by antidiscrimination law.

This condemnation of discrimination, however, is highly contingent, depending on certain empirical assumptions for much of its force. For the indictment to stand, discrimination in education, employment, and the provision of public accommodations must be sufficiently pervasive in society to diminish the life chances of its victims appreciably, to distort the functioning of the economic system enough to deprive its victims of an equal opportunity to pursue careers suited to their talents and ambitions. If, by contrast, most firms do not discriminate based on race or sex, then discrimination need not violate Rawls’s second principle of justice; conditions of fair equality of opportunity

can still obtain.⁷⁹ In this more tolerant milieu, political liberalism will provide little, if any, justification for antidiscrimination law.

This important caveat reminds us yet again that Rawls propounds a systemic, rather than an individualistic, conception of justice in general and discrimination in particular. In other words, political liberalism seeks to condemn in the first instance not the discrete discriminatory acts of individual citizens and firms, but the institutional arrangements that unfairly distribute the various benefits and burdens of social cooperation. Only insofar as individual discriminatory acts infect society's overall economy and other main institutions, engendering an unjust distribution of benefits and burdens, will they violate fair equality of opportunity and warrant a legal remedy according to political liberalism's criteria of political legitimacy.

VI. THE PRIORITY OF LIBERTY AND ANTIDISCRIMINATION LAW

If political liberalism condemns discrimination as a violation of formal and fair equality of opportunity, then it views antidiscrimination law as a remedy for infringements of the second principle of justice. And yet, Rawls subordinates the requirements of the second principle to those of the first, a ranking he dubs the "priority of liberty."⁸⁰ According to that priority, the second principle, including its requirement of fair equality of opportunity, "is always to be applied within a setting of background institutions that satisfy the requirements of the first principle."⁸¹ The priority of liberty thus "rules out exchanges ('trade-offs,' as economists say) between the basic rights and liberties covered by the first principle and the social and economic advantages regulated

⁷⁹ Cf. Shiffrin, "Race," 1650. Matt Cavanagh similarly argues that "top-down" accounts of discrimination like Rawls's condemn not the conduct itself, but the effects it tends to produce, a contingent case that lapses once those consequences no longer materialize (Matt Cavanagh, *Against Equality of Opportunity* (Oxford: Oxford University Press, 2003), 198-99).

⁸⁰ Rawls, *Political Liberalism*, 6.

⁸¹ Rawls, *Justice as Fairness*, 46.

by the difference principle [and the principle of fair equality of opportunity]”⁸² and insists that “a basic liberty can be limited or denied solely for the sake of one or more other basic liberties.”⁸³ Only after fully guaranteeing the basic rights and liberties enshrined in the first principle of justice may the government attempt to regulate social and economic inequalities according to the provisions of the second.

So defined, the priority of liberty poses a serious challenge to antidiscrimination law: if fair equality of opportunity, a requirement of the second principle, justifies antidiscrimination law, then the legislation must not infringe any of the basic rights and liberties guaranteed by the first principle. For the priority of liberty would decide all such conflicts decisively in favor of the right or liberty being infringed rather than the equality of opportunity being promoted. Before deeming antidiscrimination law legitimate according to Rawls’s political conception of justice, then, we must ensure that it does not breach the priority of liberty by abridging any of the basic rights or liberties.

Opponents of antidiscrimination law who entertain libertarian (as opposed to genuinely liberal⁸⁴) sympathies will likely seize on the priority of liberty as evidence of the incompatibility of the legislation with political liberalism. After all, these opponents might contend, antidiscrimination law inevitably infringes the liberty of the owners of private firms by preventing them from contracting freely with their fellow citizens—from hiring and terminating employees at will and restricting their clientele as they see fit. Surely, the priority of liberty forbids such violations of freedom and thus precludes antidiscrimination law, a measure for implementing the second principle of justice’s guarantee of fair equality of opportunity. To permit the government to curtail the individual liberty of private economic actors for the sake of socioeconomic equality, the

⁸² Ibid., 47.

⁸³ Rawls, *Political Liberalism*, 295.

⁸⁴ For more on the various contrasts between liberalism *sensu stricto* and libertarianism, see Samuel Freeman, “Illiberal Libertarians: Why Libertarianism Is Not a Liberal View,” *Philosophy & Public Affairs* 30 (2002): 105-51.

argument goes, would be to turn the priority of liberty—and hence political liberalism itself—on its head.

Ironically, such libertarian skeptics can take a modicum of solace from contemporary antidiscrimination law itself, which occasionally privileges the economic liberty of business owners over the imperative of socioeconomic equality. For example, an employer may discriminate indirectly on the basis of race, promulgating a generally applicable employment requirement that disproportionately disadvantages members of a certain race, so long as he can prove that the requirement is a “business necessity,” that compliance with the relevant provisions of antidiscrimination law would cause him significant economic harm.⁸⁵ Because one facet of antidiscrimination law thus contemplates considerations of economic liberty and efficiency, libertarians might wonder why others do not likewise defer to these values. In particular, why do legal prohibitions against direct discrimination recognize no similar exception for “business necessity”? If an employer may discriminate indirectly when market forces so dictate, why may he not do so directly as well, simply refusing, say, to hire black applicants when his racist clientele threaten to cease patronizing his business if he integrates his workforce? Indeed, libertarians tend to pursue precisely this line of reasoning in dismissing antidiscrimination law altogether, arguing that most applications of the legislation impose onerous, costly requirements on firms without reaping corresponding rewards for their intended beneficiaries.⁸⁶

We radically misconceive Rawls’s project in *Political Liberalism*, however, if we understand the priority of liberty to accommodate any of these libertarian concerns. For Rawls unequivocally withholds the special protection of the priority of liberty from the

⁸⁵ See, e.g., the decision of the U.S. Supreme Court in *Griggs v. Duke Power Co.*, 401 U.S. 424 (1971).

⁸⁶ See, e.g., Richard A. Posner, *The Economics of Justice* (Cambridge, Mass.: Harvard University Press, 1981), 351-63; Richard A. Epstein, *Forbidden Grounds* (Cambridge, Mass.: Harvard University Press, 1992).

absolute freedom of contract and other economic liberties underlying libertarian criticisms of antidiscrimination law. Recall that political liberalism assigns priority not to “liberty as such,” but rather to a list of distinct fundamental *liberties*, which Rawls enumerates as follows: “freedom of thought and liberty of conscience; the political liberties and freedom of association, as well as the freedoms specified by the liberty and integrity of the person; and finally, the rights and liberties covered by the rule of law.”⁸⁷ While Rawls understands the “freedoms specified by the liberty and integrity of the person” to encompass a modest right “to hold and to have the exclusive use of personal property,” he refuses to expand that right to guarantee a more robust freedom of contract.⁸⁸ Indeed, to recognize that stronger freedom would be to disavow political liberalism’s foundational commitment to socioeconomic equality, since unfettered contractual arrangements inevitably engender “excessive social economic inequalities.”⁸⁹ Unlike its libertarian counterpart, a liberal government like the one Rawls envisages reserves the right to regulate contractual relationships like those between employers and their employees in order to avert unjust distributional patterns.⁹⁰ As we have seen, because it supposes discrimination to foster such maldistribution by undermining fair equality of opportunity, political liberalism countenances antidiscrimination law—a judgment that is in no way invalidated by freedom of contract.

VII. FREEDOM OF ASSOCIATION AND ANTIDISCRIMINATION LAW

Antidiscrimination law thus does not violate the “freedoms specified by the liberty and integrity of the person,” as understood by Rawls. But do any of the other basic rights and liberties enshrined in the first principle of justice pose a more significant challenge? Political liberalism and its priority of liberty will pronounce antidiscrimination law

⁸⁷ Rawls, *Political Liberalism*, 291.

⁸⁸ *Ibid.*, 298.

⁸⁹ *Cf. ibid.*, lviii.

⁹⁰ *Cf. Freeman*, “Illiberal Libertarians,” 113-14.

illegitimate if and only if the legislation infringes one or more of these other rights or liberties in its pursuit of fair equality of opportunity. At least in the contexts of employment and the provision of public accommodations, no core application of antidiscrimination law appears to infringe “freedom of thought and liberty of conscience,” the “political liberties,” or the “rights and liberties covered by the rule of law.”⁹¹ Only “freedom of association” potentially precludes antidiscrimination law, which compels racists and sexists to associate with people they would otherwise shun.

Freedom of association would cast no doubt on the legitimacy of antidiscrimination law if Rawls simply required associations to regulate their own internal activities according to the two principles of justice. If, in other words, Rawls pursued a strategy of “congruence between the internal life and practices of voluntary groups and the public culture of liberal democracy,” insisting that “associations mirror public norms,”⁹² then associations would have to comport themselves so as always to honor the requirements of fair equality of opportunity embodied in antidiscrimination law, requirements from which freedom of association would not exempt them. But Rawls explicitly eschews a strategy of “congruence.” Instead, he distinguishes the political realm, which encompasses only the basic structure of society, from the nonpolitical realm, which includes private associations, and maintains that the principles of justice apply directly only to the former.⁹³ To be sure, we should not view the nonpolitical, associational realm as “a space . . . exempt from justice,” for the principles of justice regulate associations indirectly through the institutions of the basic structure.⁹⁴ In

⁹¹ This statement ignores those provisions of antidiscrimination law in some jurisdictions that proscribe “hate speech” and other forms of harassment, proscriptions some scholars contend violate freedom of speech. See, e.g., Eugene Volokh, “What Speech Does ‘Hostile Work Environment’ Harassment Law Restrict?,” *Georgetown Law Journal* 85 (1997): 627-48.

⁹² Nancy L. Rosenblum, “Compelled Association: Public Standing, Self-Respect, and the Dynamic of Exclusion,” in *Freedom of Association*, ed. Amy Gutmann (Princeton: Princeton University Press, 1998), 75.

⁹³ Rawls, *Political Liberalism*, 137; Rawls, *Justice as Fairness*, 10.

⁹⁴ Rawls, *Justice as Fairness*, 166.

enacting such oblique regulation, however, the government imposes requirements significantly less onerous and intrusive than those imposed by the principles of justice themselves—requirements, for example, that prevent churches from punishing heresy and apostasy but not from refusing to ordain female clergy.⁹⁵ Moreover, while the government may legitimately seek to safeguard the basic rights and liberties and promote fair equality of opportunity within associations,⁹⁶ it must not contravene those very requirements in the process. We must determine, then, whether antidiscrimination law violates the first principle’s guarantee of freedom of association by indirectly subjecting many associations to the requirements of fair equality of opportunity.

Conceptions of the Good, Freedom of Conscience, and Freedom of Association

To answer this question, in turn, we must first identify the interests protected, and the governmental actions precluded, by political liberalism’s conception of freedom of association. Like all the other guarantees of the two principles of justice, that freedom is included among the fair terms of social cooperation, the rules to which free and equal citizens would agree in conditions that situated all of them “symmetrically” and accorded none of them “bargaining advantages” over the others. To simulate such conditions, Rawls constructs a hypothetical and ahistorical choice situation—the “original position”—that provides us with a point of view from which we can evaluate proposed principles for regulating social cooperation in a democratic society of free and equal citizens. The parties in the original position, who represent the free and equal citizens who eventually will participate in the social cooperation, deliberate behind a “veil of ignorance” that occludes the particular social positions, comprehensive doctrines, and other morally arbitrary characteristics of their constituents. Because the parties lack such

⁹⁵ Cf. *ibid.*, 164.

⁹⁶ Rawls, *Political Liberalism*, 261.

knowledge, Rawls insists, any principles of justice they all endorse will be ipso facto fair.⁹⁷

According to Rawls, the parties in the original position would include a robust freedom of association among the fair terms of social cooperation and would assign it a strict priority over questions of socioeconomic equality, grounding that freedom in a broader freedom of conscience.⁹⁸ In particular, recognizing that we are free and thus have a capacity for a conception of the good, the parties would adopt principles of justice that would enable us freely to develop and to exercise that capacity. Only principles that fully secured freedom of conscience would achieve this aim, for we cannot develop and exercise our capacity for a conception of the good unless we also enjoy “the liberty to fall into error and to make mistakes”; pursuing a conception of the good requires freedom of conscience.⁹⁹ And because freedom of conscience remains a hollow promise unless we can also associate with others who share our beliefs about what is valuable in life, freedom of conscience and freedom of association “go in tandem.”¹⁰⁰ By connecting freedom of association inextricably with freedom of conscience, Rawls signifies the former’s great importance for free citizens and justifies its inclusion among the basic rights and liberties enshrined in the first principle of justice and guaranteed priority over questions of socioeconomic equality.

Purporting to analyze freedom of association in terms at least compatible with those of political liberalism,¹⁰¹ Stuart White dubs the interests grounding the freedom “integrity interests,” interests related to a person’s “freedom to shape and live

⁹⁷ See *ibid.*, 22-28.

⁹⁸ Rawls does adduce two other sets of considerations to defend freedom of conscience (see *ibid.*, 311-12, 315-24), but he considers them less relevant for freedom of association.

⁹⁹ *Ibid.*, 312-14.

¹⁰⁰ *Ibid.*, 313.

¹⁰¹ See Stuart White, “Freedom of Association and the Right To Exclude,” *Journal of Political Philosophy* 5 (1997): 377-78, 378 n.7.

authentically in accordance with a distinctive ethical personality.”¹⁰² A person’s “ethical personality,” in turn, comprises “the cluster of commitments—philosophical, religious, affective—which represent the individual’s personal response to fundamental questions of value and meaning and which thereby give his/her life its basic normative shape and direction.”¹⁰³ Because we cannot pursue these commitments without associating with other people who share them, we require freedom of association to actuate freedom of conscience.¹⁰⁴ As both White’s account and Rawls’s make clear, freedom of association seeks to protect from illegitimate governmental interference those groups in which people collectively pursue their conceptions of the good.

White’s and Rawls’s accounts of freedom of association seem to portray the groups that enjoy that freedom as rather static: having independently formed conceptions of the good, we associate with like-minded people either to disseminate our collective message to the broader public or to pursue our shared aims free from the din of opposing viewpoints. But as Seana Shiffrin rightly notes, this depiction belies the actual dynamics of many associations, whose members often seek in the company of others not approbation for beliefs they have already affirmed, but exposure to ideas they have yet to consider. We often form associations, in other words, to engage collectively in the process of *formulating* a conception of the good, not merely to report the results of our own solitary deliberations.¹⁰⁵ Despite his somewhat vague statements on the subject, Rawls recognizes this important additional function of associations, noting that “[l]iberty of conscience and freedom of association are to secure the full and informed and effective application of citizens’ powers of deliberative reason to their *forming, revising, and*

¹⁰² Ibid., 378.

¹⁰³ Ibid.

¹⁰⁴ Ibid., 380-81.

¹⁰⁵ Seana Valentine Shiffrin, “What Is Really Wrong with Compelled Association?,” *Northwestern University Law Review* 99 (1995): 865, 869.

rationaly pursuing a conception of the good over a complete life.”¹⁰⁶ We thus afford certain groups the protections of freedom of association because they provide venues for communal rational reflection as well as collective moral advocacy.

Of course, not all groups foster these valuable activities to the same degree. While some groups take the promotion of their members’ interests in developing and exercising conceptions of the good as their *raison d’être*, others pursue this aim only indirectly, viewing its achievement as a fortuitous byproduct of their other, more fundamental endeavors. “Intimate associations,” through which people attempt to secure “intimacy-related goods, such as friendship or love,” most obviously epitomize the former type of group.¹⁰⁷ Similarly, “expressive associations,” such as political parties and certain lobbying organizations, advance integrity interests by enabling their members to communicate their most fundamental beliefs publicly to their fellow citizens.¹⁰⁸ “Economic associations,” by contrast, seem to care little about helping their members to develop and to exercise conceptions of the good, instead pursuing various economic benefits almost single-mindedly.¹⁰⁹ As this typology suggests, different kinds of associations further the integrity interests of their members to different degrees and thus merit different levels of protection from governmental interference under freedom of association.

Associations differ not only in terms of the purposes they avow, but also in terms of the relationships they foster. The members of “primary associations,” such as families and groups of friends, engage in highly intimate and sustained relationships, whereas participants in “secondary associations” meet regularly in larger, less familiar gatherings. The members of “tertiary associations” form even more impersonal relationships, which

¹⁰⁶ Rawls, *Political Liberalism*, 335 (emphasis added).

¹⁰⁷ White, “Right to Exclude,” 390.

¹⁰⁸ *Ibid.*, 381.

¹⁰⁹ *Ibid.*, 379.

tend to be mediated through remote, centralized organizations.¹¹⁰ But unless it seeks to realize “intimacy-related goods,” an association can assume almost any form to pursue a particular goal—engaging in expressive activities, for example, while sustaining only “tertiary” relationships. According to political liberalism, a group should enjoy the protections of freedom of association so long as it espouses goals or fosters relationships that help its members to develop and to exercise their capacities for a conception of the good.

Foremost among these protections is an association’s right to exclude people who reject its fundamental tenets and purposes, a right that potentially thwarts much antidiscrimination law. If an association could not determine its own membership, if the government might compel it to admit even those people intent on subverting its purposes, it would lose its cohesive, communal character and, consequently, its ability to advance the integrity interests of its other members.¹¹¹ To be sure, an association sometimes excludes people not to further its avowed purposes, but to realize other, trivial objectives or to regulate some of its more peripheral activities.¹¹² While it is often difficult to distinguish necessary exclusions from gratuitous ones, freedom of association seems to insulate from governmental regulation only the former, “purpose-protecting” ones—only those that attempt to preserve an association’s communal integrity.¹¹³ Such essential exclusions, moreover, often will be “categorical”—that is, they will “exclude an individual from membership because he/she belongs to a community defined by possession of some shared characteristic such as race, gender, sexuality, or religion.”¹¹⁴

¹¹⁰ Amy Gutmann, “Freedom of Association: An Introductory Essay,” in *Freedom of Association*, ed. Amy Gutmann (Princeton: Princeton University Press, 1998), 10.

¹¹¹ White, “Right to Exclude,” 373, 381. Nor is it obvious that the government is even capable of thwarting associations that are determined to discriminate. See Larry Alexander, “What Is Freedom of Association and What Is Its Denial?,” *Social Philosophy and Policy* 25 (2008): 8.

¹¹² See Kent Greenawalt, “Freedom of Association and Religious Association,” in *Freedom of Association*, ed. Amy Gutmann (Princeton: Princeton University Press, 1998), 115-20.

¹¹³ White, “Right to Exclude,” 380.

¹¹⁴ *Ibid.*, 374-75.

Insofar as its conception of freedom of association protects these categorical exclusions from governmental regulation, political liberalism will deem antidiscrimination law illegitimate.

Legitimate and Illegitimate Applications of Antidiscrimination Law

The application of antidiscrimination law is most obviously problematic in the cases of intimate and expressive associations. The former seem to enjoy an absolute right to exclude people on any grounds whatsoever, since the activity of forming intimate relationships is so fundamental to developing and exercising one's capacity for a conception of the good. But which associations are truly intimate? White proposes two plausible criteria: "Firstly, there must be a strong and mutual familiarity, ordinarily grounded in regular, intensive 'face to face' interaction, between more or less all organization members. . . . Secondly, the pursuit and enjoyment of intimacy-related goods, such as friendship or love, must be the organization's primary associative purpose."¹¹⁵ Though relatively few associations will satisfy these rather demanding criteria, many private social clubs—even some with fairly large memberships—will, entitling them to the categorical right to exclude bestowed by political liberalism's conception of freedom of association. The antidiscrimination regimes of most jurisdictions recognize this right by declining to impose their requirements on smaller, more intimate associations.

Nor would political liberalism countenance the application of antidiscrimination law to expressive associations—at least when the legislation would oblige an association to admit people who reject the views it seeks to propagate.¹¹⁶ By compelling, say, a white supremacist organization to admit black people, the government would thoroughly compromise the organization's efforts to disseminate its message. However morally retrograde this message might be, political liberalism's conception of freedom of

¹¹⁵ Ibid., 390.

¹¹⁶ Ibid., 381.

association safeguards a group's right to communicate it and to exclude anyone who repudiates it. Expressive associations, moreover, enjoy this categorical right to exclude regardless of the nature of the interactions among their members—that is, regardless of whether they are primary, secondary, or tertiary expressive associations—since their purposes directly implicate freedom of conscience and their members' capacities for a conception of the good. Freedom of association thus exempts at least those expressive associations that advocate racist and sexist beliefs from the requirements of antidiscrimination law.

If political liberalism obviously precludes the application of antidiscrimination law to many intimate and expressive associations, it seems just as obviously to subject most economic associations to the legislation. Indeed, most economic firms appear not even to qualify as associations for the purposes of freedom of association. A firm might attempt to assert a putative right to exclude in two contexts: when interacting with its prospective or current employees or when providing goods and services to its customers. In neither context, however, do most economic associations seem to enjoy freedom of association and the attendant right to exclude, since they attempt simply to turn a profit rather than to further anyone's interest in developing and exercising a capacity for a conception of the good. By excluding a person in either context, moreover, an economic firm harms his "opportunity interests," depriving him of "fair access to income and wealth, and to other goods, like education and employment, which have clear instrumental value in enabling the individual to realise his/her ethical personality."¹¹⁷ Were it to accord economic firms the protections of freedom of association, the government would have to tolerate such violations of people's "opportunity interests" and, consequently, the significant socioeconomic inequalities that would likely ensue. As a

¹¹⁷ Ibid., 378.

theory devoted to guaranteeing socioeconomic equality, political liberalism thus seems to have no compunction about applying antidiscrimination law to economic associations.

But how accurate is the preceding portrayal of the purposes and activities of economic associations? Do no firms take as one of their primary purposes the promotion of their members' interests in formulating and pursuing conceptions of the good and consequently merit the protections of freedom of association and the attendant right to exclude? At least in the context of employment, some small businesses attempt through their decisions regarding hiring, pay, and so on not simply to turn a profit, but also to realize certain "intimacy related goods" like friendship and the community of engaging in a cooperative enterprise, as well as to provide their members a venue in which to form and pursue conceptions of the good. Indeed, many people view meaningful and rewarding employment as both a constituent element of the good life and a uniquely conducive setting for its partial realization¹¹⁸—a richer conception of work than a reductive one that focuses exclusively on its pecuniary benefits. We thus should conceive of at least some small businesses not only as economic associations, but also as intimate associations, associations on which political liberalism would confer the protections of freedom of association, including a robust right to exclude. In preventing the members of these economic-*cum*-intimate associations from selecting their co-workers according to their preferred criteria, the government would compromise the social relationships they had sought to cultivate, impeding the development and exercise of their capacities for a conception of the good. It is difficult to see on what grounds political liberalism may countenance applications of antidiscrimination law that have these deleterious effects.¹¹⁹

¹¹⁸ Shiffrin, "Race," 1666-67. In support of these assertions, Shiffrin cites (*ibid.*, 1666 n.71), among other works, Cynthia Estlund, *Working Together: How Workplace Bonds Strengthen a Diverse Democracy* (Oxford: Oxford University Press, 2003); and Vicki Shultz, "Life's Work," *Columbia Law Review* 100 (2000): 1881-1964.

¹¹⁹ Title VII of the U.S. Civil Rights Act of 1964 exempts businesses with fewer than fifteen employees even from its provisions governing employment discrimination (§ 701(b)). As the next paragraph argues,

Indeed, political liberalism's conception of freedom of association sweeps more broadly than we initially might suppose, extending its protections even to many large economic firms. For just like their more intimate counterparts, these firms furnish their members with opportunities to form and pursue conceptions of the good. To be sure, large economic firms will enjoy the freedom of association and the attendant right to exclude in fewer contexts, since only sometimes must such firms categorically exclude entire classes of people to cultivate their members' capacities for a conception of the good. Only with great disingenuousness, for instance, could the male employees in a large corporation's marketing division claim that they had to exclude women from the corporation's mailroom to preserve the relationships through which they pursued their conceptions of the good. Nonetheless, their claim might be more plausible were they attempting to exclude women from their own division, a relatively intimate context that might resemble a small business more than a diffuse enterprise. It thus seems that political liberalism must formulate institutional principles that recognize at least some assertions of freedom of association and the attendant right to exclude by larger economic associations—namely, those assertions that genuinely help people to formulate and pursue their conceptions of the good.

We might attempt to resist this implication of political liberalism's conception of freedom of association in two ways. First, we might note that, even if an economic association provides its members with "intimacy related goods" and cultivates their capacities for a conception of the good, it always seeks above all else to realize certain economic benefits, remaining a business rather than reconstituting itself as a social club.¹²⁰ Moreover, economic associations necessarily "operate within a highly competitive marketplace and have a fairly focused singular purpose whose pursuit is

however, political liberalism's conception of freedom of association requires exemptions not only for such small businesses, but also for many larger ones.

¹²⁰ Cf. White, "Right to Exclude," 390.

largely guided by this competitive context and aim of profitable operation,” a milieu largely inhospitable to the kinds of intimate relationships through which people formulate and pursue their conceptions of the good.¹²¹ Only if it extricates itself from economic markets and espouses as its *primary* purpose the promotion of intimate relationships, the argument goes, will an association effectively foster its members’ capacities for a conception of the good and thus merit the safeguards of freedom of association, including the right to exclude.

On initial inspection, it appears that Rawls espouses precisely this conception of freedom of association, contrasting economic associations—irrespective of their other purposes—with noneconomic ones and according the protections of freedom of association only to the former. For instance, according to Kevin Kordana and David Blankfein Tabachnick, Rawls conceives of freedom of association not as one single, monolithic entitlement, but rather as a “complex right” that “conjoins a set of personal liberties (e.g., the freedom of assembly and the freedom to form intimate relationships) with a set of economically oriented liberties (i.e., the rights of private ordering).”¹²² Only the former, personal liberties, they argue, are enshrined in Rawls’s first principle of justice and granted priority over questions of socioeconomic equality. For according to what Kordana and Tabachnick dub the “high Rawlsian position,” “all robust economic constructions and baselines . . . are properly understood, for Rawls, as second-principle matters”—constructions and baselines that constrain the “economic component” of freedom of association.¹²³ Were Rawls instead to entrench even this component in the first principle of justice and thereby to insulate many economic associations from the demands of the second, they worry, he would effectively abjure his commitment to

¹²¹ Shiffrin, “Compelled Association,” 877.

¹²² Kevin A. Kordana and David H. Blankfein Tabachnick, “The Rawlsian View of Private Ordering,” *Social Philosophy and Policy* 25 (2008): 290.

¹²³ *Ibid.*, 292.

socioeconomic equality, instituting an economic system so impervious to governmental regulation that it could never honor fair equality of opportunity or the “difference principle.”¹²⁴ Unless it is willing to embrace the libertarian freedom of contract, political liberalism must apparently deny all economic firms the protections of freedom of association.

Were we to endorse this bifurcation of freedom of association into a protected, “personal” component and a vulnerable, “economic” one, however, we would effectively emasculate the liberty, subverting Rawls’s rationale for enshrining it in the first principle of justice in the first place. For many—perhaps most—citizens of contemporary western democracies develop and exercise their capacities for a conception of the good not just in groups chiefly and explicitly devoted to that purpose, but also in groups that realize it only incidentally, as a byproduct of their other, principal endeavors, such as profit making. To be sure, some people attend churches and other houses of worship, but many others form and pursue their conceptions of the good in social clubs, expressive associations, and sometimes even businesses and other economic enterprises.¹²⁵ *Pace* Kordana and Tabachnick, then, we cannot in practice distinguish between the economic and non-economic activities of a particular association; any demarcation of the two kinds of activities will necessarily be “permeable” and “artificial.”¹²⁶ Given this nearly seamless integration of commercial relationships and more intimate ones, “[a]ny governmental interference with freedom of any form of association, including commercial association, will perforce result in interference with the other forms of association.”¹²⁷ To deny economic firms the protections of freedom of association categorically would be to

¹²⁴ *Ibid.*, 293-94.

¹²⁵ As numerous sociologists have noted, Americans have increasingly turned to the workplace—and away from civic societies, houses of worship, and neighborhood associations—to form friendships and other personal relationships. See, e.g., Robert Putnam, *Bowling Alone: The Collapse and Revival of American Community* (New York: Simon and Schuster, 2000), 85-92.

¹²⁶ Alexander, “Freedom of Association,” 13.

¹²⁷ *Ibid.*, 14.

deprive many people of important venues for forming and pursuing conceptions of the good and thus to thwart one of political liberalism's cardinal aims.

According to the second argument for withholding the protections of freedom of association from all economic associations, while such associations often help their members to form conceptions of the good, they must harm the opportunity interests of those people whom they categorically exclude to do so—a significant harm that trumps freedom of association. White seems to mount a version of this defense, contending that “individuals have specific rights against categorical exclusion which protect their opportunity interests.”¹²⁸ White founds this claim on an “egalitarian” conception of justice that holds that “people ought to have equal access to goods like income and wealth,”¹²⁹ a requirement that seems to comport with at least the spirit of Rawls's notion of fair equality of opportunity. This conception of justice, White claims, deems categorical exclusions that unfairly harm people's opportunity interests “presumptively unjust.”¹³⁰ Therefore, White concludes, antidiscrimination law should apply to all economic associations—even those that might also count as intimate associations—as a requirement of egalitarian justice.

Though we might find this response intuitively appealing, it is unavailable to Rawls. In particular, it threatens to upend the priority of liberty, subordinating a guarantee of the first principle of justice to a requirement of fair equality of opportunity. If some economic associations actually do enjoy the protections of freedom of association because they significantly promote their members' interests in formulating and pursuing conceptions of the good, then, according to political liberalism, they do not automatically forfeit those protections if they happen to violate fair equality of opportunity. Instead, the priority of liberty requires the government to defer in its pursuit of fair equality of

¹²⁸ White, “Right to Exclude,” 380.

¹²⁹ *Ibid.*, 383.

¹³⁰ *Ibid.*

opportunity to the basic rights and liberties, including freedom of association, a requirement that insulates the employment decisions of at least some economic associations from the strictures of antidiscrimination law.

Does political liberalism also impede the application of antidiscrimination law to economic associations in the context of the provision of public accommodations? May an economic association ever exclude customers as well as job applicants on the basis of their race or sex? Political liberalism's conception of freedom of association seems to afford economic associations much less discretion in this context than in that of employment, since customers are not, strictly speaking, members of the associations. Nevertheless, freedom of association would permit economic associations to discriminate against potential customers if necessary to preserve their associational integrity and to foster their members' capacities for a conception of the good. If, for example, an association's employees had to engage in frequent, sustained, and intimate interactions with its customers—if, that is, the customers were *de facto* members of the association—then the association might invoke freedom of association to exclude those potential customers whose mere presence would compromise its intimacy-related purposes. As in employment, then, political liberalism's conception of freedom of association permits some economic associations to discriminate in the provision of public accommodations, forestalling some core applications of antidiscrimination law.

Not all discriminators, of course, discriminate to further their conceptions of the good. Rather, much discrimination based on race or sex stems not from any comprehensive philosophical doctrine, but from brute economic calculation. Many employers, for example, use race or sex as a proxy for other personal traits that they deem relevant to their firms' respective enterprises, while others discriminate to appease their customers' prejudice rather than to indulge their own. Because the perpetrators of these

and other similarly motivated instances of discrimination seek to further no particular conception of the good, they may not invoke freedom of association to exempt themselves from the requirements of antidiscrimination law. Ironically, then, political liberalism ends up protecting only the most deplorable discriminatory acts, ones motivated by outright racism or sexism rather than prosaic economic ambition.

With its priority of liberty and guarantee of freedom of association, Rawls's political conception of justice legitimates some important applications of antidiscrimination law but condemns many others. Most strikingly, to reconcile antidiscrimination law with political liberalism, we first would have to qualify the legislation as it applies to many economic associations, countenancing those discriminatory acts through which such associations attempt to promote their members' conceptions of the good. As currently constituted, the antidiscrimination regimes of contemporary western liberal democracies contravene political liberalism's criteria of political legitimacy.

To be fair, Rawls himself never unequivocally delineates the boundaries of freedom of association. He thus could conceivably disavow the foregoing analysis of that freedom, insisting, for example, that no economic firm—no matter how intimate—enjoys its protections and that antidiscrimination law consequently applies straightforwardly to the vast majority of associations. By taking such an uncompromising approach to economic associations, however, Rawls would attenuate the connection he draws between freedom of association and the capacity for a conception of the good, exposing many of the venues in which people exercise that capacity to intrusive governmental interference. Insofar as it grounds freedom of association in freedom of conscience and the capacity for a conception of the good, political liberalism will insulate many economic associations from such interference and forestall many applications of antidiscrimination law.

The tension between political liberalism and antidiscrimination law thus ultimately owes to the theory's antiperfectionism—in particular, its antiperfectionist conception of freedom of association. Because it insists on maintaining scrupulous neutrality among competing conceptions of the good life, political liberalism can espouse only the most emaciated account of the interests underlying freedom of association. In particular, because political liberalism construes freedom of association merely as a means for developing and pursuing one's conception of the good, whatever the conception's content, it cannot distinguish between ennobling and corrupting acts of association. To draw such distinctions, we would have to espouse a perfectionist understanding of freedom of association and the interests it protects, unabashedly advertent to a conception of the good.

VIII. BEYOND THE TWO PRINCIPLES OF JUSTICE?

Thus far, this chapter has attempted to take political liberalism on its own terms rather than to reconstruct the theory better to accommodate antidiscrimination law. For example, having demonstrated the failure of the first principle of justice to justify antidiscrimination law and noted Rawls's various statements construing discrimination as a violation of the second principle's guarantee of fair equality of opportunity, this chapter has assumed that political liberalism's condemnation of discrimination and justification for antidiscrimination law must be grounded in the second principle. It thus has not considered how we might reformulate the first principle to incorporate a right to nondiscrimination, a revision that would altogether avoid the problems presented by the priority of liberty. Nor has it tapped other, underdeveloped elements of Rawls's theory, such as the individualistic "natural duties," which might provide a nonsystemic, but nevertheless antiperfectionist, account of discrimination and antidiscrimination law. We

thus might wonder whether we can revise political liberalism to legitimate all applications of antidiscrimination law while preserving the theory's core features.

Perhaps the least extensive revision would simply relocate political liberalism's nondiscrimination norm from the second principle of justice to the first. Seana Shiffrin, for instance, would include a "prohibition on arbitrary discriminatory treatment" among the basic rights and liberties guaranteed by the first principle, a prohibition that she believes the parties in Rawls's original position would endorse to secure their "equal status."¹³¹ Noting the "obvious compatibility of anti-discrimination principles with the Rawlsian framework," moreover, Shiffrin contends that this amendment raises no "theoretical difficulty" for Rawls's conception of justice.¹³² Were it to adopt Shiffrin's proposal, political liberalism would indeed remove one of the more significant impediments to antidiscrimination law, grounding its condemnation of discrimination in the first principle of justice rather than the second and thus circumventing the priority of liberty.

Pace Shiffrin, however, it is by no means obvious that the parties in the original position would enshrine a categorical prohibition against discrimination in the first principle of justice. For the parties would assess the prohibition, like any other proposed rule of social regulation, according to "maximin" reasoning, the reasoning of the members of the least well-off socioeconomic class. Given the possibility that some forms of discrimination, in some empirical conditions, might actually maximize the well-being of the least well-off, the first principle could incorporate not Shiffrin's categorical prohibition against discrimination, but, at most, only a more qualified one—namely, one that condoned any discrimination that would maximize the well-being of the least well-

¹³¹ Shiffrin, "Race," 1660. For a similar proposal, see Susan Moller Okin, *Justice, Gender, and the Family* (New York: Basic Books, 1989), ch. 5.

¹³² Shiffrin, "Race," 1661, 1660.

off.¹³³ And regardless of the strength of the antidiscrimination norm it espoused, political liberalism would still have to resolve the inevitable conflicts between that norm and the other guarantees of the first principle—conflicts Shiffrin herself foresees.¹³⁴ Most notably, it would continue to confront the very conflicts between nondiscrimination and freedom of association considered in this chapter.

How would a reconceived version of political liberalism like Shiffrin's adjudicate these conflicts? While it seems that the theory would at least occasionally subordinate freedom of association to nondiscrimination, thereby legitimating considerably more applications of antidiscrimination law than it could by treating discrimination as a violation of fair equality of opportunity, it would still have to tolerate many discriminatory economic firms. For if such firms truly promote the interests underlying freedom of association, if they truly foster their members' capacities for a conception of the good, then political liberalism cannot thwart them without subverting its own *raison d'être*: the development of the two moral powers in which citizens' personal autonomy consists. For a similar reason, political liberalism cannot simply dissolve the priority of liberty and privilege fair equality of opportunity over freedom of association without radically altering its normative underpinnings—indeed, without abjuring its liberalism altogether.¹³⁵

While Shiffrin's proposal would alter the precise normative basis of political liberalism's case against discrimination, it would nevertheless leave that case's systemic structure intact. For just like the second principle of justice, the first purports directly to govern the arrangement of society's main institutions rather than the actions of individual citizens. A general right to nondiscrimination would thus require the basic structure to

¹³³ Arneson, "Against Rawlsian Equality of Opportunity," 100-02.

¹³⁴ Shiffrin, "Race," 1661.

¹³⁵ See generally Robert S. Taylor, "Rawls's Defense of the Priority of Liberty: A Kantian Reconstruction," *Philosophy & Public Affairs* 31 (2003): 246-71.

treat all citizens equally, regardless of their race or sex; it would not directly prohibit individual persons from discriminating on those bases. To be sure, the government would most likely have to enact such an individual prohibition, in the form of antidiscrimination law, in order to ensure that the basic structure complied with any nondiscrimination principle. But Shiffrin's proposed right to nondiscrimination would remain systemic in form, shaping institutional structure in the first instance and only derivatively restricting individual conduct.

Unlike the two principles of justice, however, other features of Rawls's theory are not systemic. Most notably, in his initial exposition of his theory of justice, Rawls identifies certain "natural duties" that directly govern not institutional arrangements, but individual conduct. Such duties include "the duty of helping another when he is in need or jeopardy, provided that one can do so without excessive risk or loss to oneself; the duty not to harm or injure another; and the duty not to cause unnecessary suffering."¹³⁶ More relevant for the issue of antidiscrimination law are the natural duties of justice and "mutual respect." The first "requires us to support and to comply with just institutions that exist and apply to us. It also constrains us to further just arrangements not yet established."¹³⁷ The second is "the duty to show a person the respect which is due to him as a moral being, that is, as a being with a sense of justice and a conception of the good."¹³⁸ Rather than derive an account of discrimination and antidiscrimination law from either of the principles of justice, as Rawls himself and most of his followers explicitly do, we might look to these two natural duties to provide such an account. And perhaps that alternative basis for condemning discrimination and justifying antidiscrimination law could avoid some of the problems identified in this chapter.

¹³⁶ Rawls, *Theory of Justice*, 98.

¹³⁷ *Ibid.*, 99.

¹³⁸ *Ibid.*, 297.

Foremost among those problems was the systemic nature of Rawls's account of discrimination. By grounding that account in the systemic value of socioeconomic equality, Rawls placed it in direct conflict with his theory's antiperfectionism, which requires the government to abide certain kinds of individual conduct and thus impedes the realization of some systemic goals. An account of discrimination grounded in the individualistic natural duties, by contrast, would seem to avoid this conflict. Because such an account is independent of the second principle of justice, it is not subject to the priority of liberty. And because it is independent of the first principle, it need not accommodate other basic rights and liberties constraining the design of institutions, but rather can apply directly to individual persons, regardless of the particular associations they might form or activities in which they might engage. Not only are the natural duties independent of the two principles; they are individualistic rather than systemic. That is, they "have no necessary connection with institutions or social practices" and "hold between persons irrespective of their institutional relationships."¹³⁹ Therefore, any account of discrimination grounded in the natural duties would be nonsystemic, directly forbidding individual persons to discriminate.

The natural duties become an even more plausible source within Rawls's theory for an account of discrimination once we realize that they also ground his theory's account of criminal law. Although Rawls does not focus on criminal law and punishment, his limited discussion of those institutions makes clear that they are grounded not in the principles of justice, but in the natural duties. Thus, Rawls asserts that "the purpose of the criminal law is to uphold basic natural duties, those which forbid us to injure other persons in their life and limb, or to deprive them of their liberty and property."¹⁴⁰ If

¹³⁹ Ibid., 98, 99.

¹⁴⁰ Ibid., 276.

criminal law enforces duties that apply directly to persons rather than institutions, then so, too, might antidiscrimination law. The two kinds of legislation seem to go hand-in-hand.

There are a number of problems, however, with locating Rawls's account of discrimination in the natural duties rather than the principles of justice. First, criminal law cannot evade the requirements of the two principles of justice simply because it enforces the individualistic natural duties rather than some systemic value like socioeconomic equality. For criminal law is undoubtedly one of the institutions composing the basic structure, and as such, it must comply with the requirements governing those institutions—namely, the two principles. In particular, it must comply with the basic rights and liberties enshrined in the first principle—most notably, the rule of law, but also freedom of association. To be sure, given the nature of the conduct regulated by criminal law, the legislation will rarely infringe those rights and liberties, properly construed. But in the marginal cases where it does, it will have to yield, just like any other institution of the basic structure. Similarly, even if we construe antidiscrimination law as enforcing certain natural duties rather than securing socioeconomic equality, it will still have to observe the basic rights and liberties. And unlike criminal law, antidiscrimination law will tend to violate those rights and liberties relatively frequently—both because discriminatory conduct tends to be less severe than most criminal conduct and because, as this chapter has shown, antidiscrimination law often restricts activities that implicate political liberalism's core commitment to autonomy. Thus, despite its individualistic nature, an account of antidiscrimination law rooted in the natural duties will still conflict with political liberalism's antiperfectionism. It is ultimately political liberalism's antiperfectionist commitment to personal autonomy, not just its systemic account of socioeconomic equality, that thwarts much antidiscrimination law.

Nor is it at all clear that the natural duties actually condemn discrimination based on race or sex—or at least many of the instances of the conduct proscribed by contemporary antidiscrimination law. Although Rawls never provides an exhaustive list of the natural duties, the ones identified above seem to furnish the most plausible grounds for a duty not to discriminate. And yet, none of them seem sufficiently robust to justify antidiscrimination law. For example, Rawls seems to conceive of “the duty not to harm or injure another” as a prohibition against causing pain or physical injury to others, not against depriving others of economic benefits like employment or access to public accommodations. And the natural duty of justice imposes not a general duty to treat others fairly, but rather a more limited duty “to support and to comply with just institutions.” To be sure, even this more limited duty will require us to obey antidiscrimination law (insofar as the legislation comports with the two principles of justice). But while the natural duty of justice might demand *compliance* with antidiscrimination law once instituted, it is question-begging to invoke the duty also to justify the *adoption* of the legislation in the first place. Finally, it seems we can derive an account of discrimination from the duty of “mutual respect” only if we adopt a thicker conception of “respect” than Rawls does—indeed, the kind of conception that political liberalism’s antiperfectionism appears to preclude. In short, whereas Rawls’s natural duties apply to persons simply qua persons, the kinds of nondiscrimination norms embodied in contemporary antidiscrimination law govern their conduct only insofar as they occupy certain kinds of roles in certain kinds of social institutions. Given political liberalism’s rigid bifurcation between norms that apply universally, which are embodied in the natural duties, and norms that presuppose a particular institutional context, which are embodied in the two principles of justice, it seems more faithful to the theory to attempt to found antidiscrimination law on the latter set of norms—a conclusion

supported by Rawls's own statements linking antidiscrimination law to the second principle's guarantee of fair equality of opportunity.

Rawls made most of those statements in his later work—that is, after he attempted to dissociate his theory of justice from any comprehensive conception of the good. By contrast, his account of the natural duties appears only in his earlier work; the index to *Political Liberalism*, for instance, contains no reference to the natural duties. We thus might legitimately question whether the natural duties survive the transition from *A Theory of Justice* to *Political Liberalism*, whether they accord equally with Rawls's political conception of justice and his liberal comprehensive conception of the good. There is good reason to think that the natural duties are indeed in tension with political liberalism. For in expounding criminal punishment, whose purpose is to enforce the natural duties, Rawls claims that “a propensity to commit such acts [i.e., violations of the natural duties] is a mark of bad character, and in a just society legal punishments will only fall upon those who display these faults.”¹⁴¹ But political liberalism, of course, forbids the government to take account of people's “bad character[s]” and “faults”; rather, the antiperfectionist liberal principle of legitimacy requires the government to adduce some neutral justification, one rooted in public reason rather than a comprehensive conception of the good, for restricting people's conduct. It seems that Rawls can incorporate the requirements imposed by the natural duties into political liberalism in one of two ways. First, he can espouse a thinner conception of the natural duties, one that prescind from people's characters and faults and that consequently complies with the dictates of public reason. Alternatively, and more radically, he can deny that any duties apply directly to persons and instead derive the individualistic requirements composing criminal law indirectly from the first principle of justice. Given the absence of the natural

¹⁴¹ Ibid., 277.

duties from his exposition of political liberalism, there is good reason to think that Rawls would choose the latter, institutionalist alternative.¹⁴² But whichever one he would espouse, the natural duties would be unlikely to sustain an account of antidiscrimination law any more robust than the one grounded in fair equality of opportunity. Once again, Rawls's antiperfectionism severely constrains the pursuit of his other objectives, including nondiscrimination.

IX. CONCLUSION

Political liberalism forestalls many core applications of antidiscrimination law. Subsuming discrete discriminatory acts under a broader social practice of exclusion and disadvantage, the theory condemns discrimination based on race or sex not for its intrinsic immorality, but rather for the unjust systemic effects it tends to produce. More specifically, according to political liberalism, discrimination tends to erode the institutional preconditions of fair equality of opportunity, a requirement of the second principle of justice. But political liberalism categorically subordinates the requirements of the second principle to those of the first, including freedom of association. While, on initial inspection, economic firms appear not to enjoy this freedom, many of them actually foster their members' capacities for a conception of the good—the precise interest that political liberalism's conception of freedom of association purports to protect. To grant freedom of association priority over questions of socioeconomic equality, then, is to insulate many economic firms from the requirements of fair equality of opportunity, foremost among which is the prohibition against discrimination based on race or sex.

The priority of freedom of association and the other basic rights and liberties over questions of socioeconomic equality is an ineliminable feature of political liberalism—in particular, of the theory's antiperfectionism. It reflects political liberalism's commitment

¹⁴² Statements even in *A Theory of Justice* support this reading. See, e.g., *ibid.*, 212 (“[T]he principles justifying [penal] sanctions can be derived from the principle of liberty.”).

to autonomy, the theory's willingness to insulate certain kinds of individual conduct from governmental interference as a means of ensuring that the government heeds the requirements of public reason. But this tolerance comes at a significant cost to the theory's other objectives. Most notably, the antiperfectionism of the first principle of justice thwarts political liberalism's other systemic goals, including fair equality of opportunity and the duties of nondiscrimination it ostensibly grounds. It is its simultaneous adoption of an antiperfectionist account of individual liberty and systemic account of discrimination that causes political liberalism to condemn many applications of antidiscrimination law.

Chapter 3

The Harm Principle

I. INDIVIDUALISTIC WRONGS AND ANTIPERFECTIONISM

The previous chapter considered an antiperfectionist doctrine of political legitimacy that also propounds a systemic account of discrimination based on race or sex and concluded that such a doctrine fails to justify certain core applications of contemporary antidiscrimination law. Because the pursuit of systemic goals inevitably infringes personal autonomy, any antiperfectionist doctrine of political legitimacy that espouses such goals will have to adopt a bifurcated normative system, enshrining certain individual rights and liberties along with the systemic goals and ultimately assigning the former priority over the latter. Consequently, if an antiperfectionist doctrine of political legitimacy conceives of discrimination as a systemic injustice, it will end up thwarting those applications of antidiscrimination law that impinge on basic rights and liberties, understood in antiperfectionist terms. As the previous chapter's examination of Rawls's theory of political liberalism demonstrated, such impingement is likely to occur rather frequently, given antiperfectionism's capacious conception of the rights and liberties needed to safeguard personal autonomy. To espouse antiperfectionism and a systemic account of discrimination simultaneously is thus to invite a conflict between the ideals of personal autonomy and nondiscrimination, a conflict that contemporary liberal theories, at least, will resolve in favor of the former.

Contemporary liberals might attempt to avoid this conflict and legitimate the full extent of antidiscrimination law by jettisoning systemic accounts of discrimination for an individualistic one. If a liberal theory must privilege its individualistic, autonomy-based commitments over any systemic ones it might espouse, then perhaps it can reconcile the

values of autonomy and nondiscrimination simply by construing the latter as one implication of the former. If, in other words, a liberal theory conceives of each discriminatory act as a discrete violation of autonomy rather than as an indirect instantiation of a broader, systemic injustice, then perhaps it will forestall fewer applications of antidiscrimination law, which, on this view, serves not to promote some systemic conception of distributive justice, but rather to secure every person's autonomy-based rights. Put simply, an individualistic account of discrimination appears to sit more comfortably with liberalism's individualistic commitment to personal autonomy.

It seems, moreover, that an individualistic account of discrimination can remain antiperfectionist. For if we focus on the effects individual discriminatory acts have for their victims, then we can condemn discrete discriminatory acts without embroiling ourselves in intractable moral disputes about notions of individual character, virtue, and so on— notions inextricably bound up with controversial questions about the good life. A doctrine of political legitimacy that condemns individual instances of discrimination and licenses individual applications of antidiscrimination law while nevertheless remaining thoroughly antiperfectionist thus seems not only plausible, but also highly desirable— especially in societies as fraught with moral discord as ours is.

Considered by many contemporary liberal political philosophers to be the touchstone of legitimate governmental action, the so-called harm principle presents itself as just such a doctrine. Indeed, John Stuart Mill, perhaps the harm principle's foremost proponent, first presented the doctrine partly as a means of extricating society from moral controversies. Lamenting the fact that people "have occupied themselves rather in inquiring what things society ought to like or dislike, than in questioning whether its

likings or dislikings should be a law to individuals,”¹ Mill hoped the harm principle would help to shift their focus from the former question to the latter, from an assessment of the morality of individual behavior to one of the legitimacy of governmental regulation.² By establishing the prevention of harm as the overriding criterion of political legitimacy, the harm principle promises to favor such a shift, since it seems we need assess neither the morality of individual behavior nor the character of the persons who engage in it to pronounce such behavior harmful and thus susceptible to governmental regulation. On the contrary, because harmfulness is a property that inheres in the effects of actions rather than in the motives of actors, the harm principle seems to render questions of individual morality almost irrelevant to determining the legitimacy of governmental regulation. On this view, we should look to the harm principle not “to identify conduct that is morally objectionable or to explain why it is so,” but “to identify those actions that are legitimate candidates for a legal prohibition, whether they are independently objectionable or not.”³ A distinctly political norm rather than the cornerstone of a comprehensive moral theory, the harm principle coins a normative idiom for assessing the legitimacy of governmental regulation in which we all are fluent, regardless of any other views we might have about the wrongfulness of particular forms of individual behavior.

This understanding of the harm principle associates the doctrine rather closely with the liberal ideal of public reason, as explicated in the Introduction. Recall that, according to that ideal, we must seek to justify the exercise of political power to our fellow citizens in terms they all can accept without having to disavow their most fundamental moral beliefs. Were governmental regulation to be justified on more partisan

¹ John Stuart Mill, *On Liberty and Other Essays*, ed. John Gray (Oxford: Oxford University Press, 1998), 11.

² A similar concern motivates H.L.A. Hart’s famous account of the harm principle in *Law, Liberty, and Morality* (Stanford: Stanford University Press, 1963).

³ Colin Bird, “Harm Versus Sovereignty: A Reply to Ripstein,” *Philosophy & Public Affairs* 35 (2007): 192.

grounds, the argument went, many citizens would have to choose between recognizing the government's authority and compromising their autonomy, the capacity that, according to liberalism, most defines their personhood. The harm principle, by contrast, purports to reconcile governmental authority and individual autonomy by subjecting certain objectionable forms of individual behavior to governmental regulation without having to advert to controversial notions of the good, thereby permitting broad agreement about the legitimate purview of political power even in the face of equally pervasive disagreement about the wrongfulness of various forms of individual behavior. In particular, the harm principle suggests an individualistic account of discrimination that justifies antidiscrimination law as a measure not for reforming the characters of those who discriminate, but rather for addressing the adverse effects discrimination has for its victims. In maintaining this focus on the victims rather than on the perpetrators of discrimination, the harm principle seeks simultaneously to adopt an individualistic perspective, to embrace antiperfectionism, and to heed the strictures of public reason.

The compatibility of antiperfectionist liberalism and antidiscrimination law depends on the viability of this balancing act. If, on closer inspection, an individualistic, antiperfectionist account of discrimination cannot justify all the core applications of contemporary antidiscrimination law, then it becomes difficult to see how contemporary liberalism more generally can do so. The harm principle seems to give liberalism the best chance of averting this undesirable outcome, for the concept of harm transcends competing conceptions of the good life; all conceptions agree that the government should at least prevent its citizens from harming one another. The failure of the harm principle to subject all discriminatory acts to governmental regulation would thus portend similar failure for other likeminded liberal doctrines of political legitimacy, which employ criteria no more ecumenical than that of harm to others.

This chapter argues that an antiperfectionist version of the harm principle does indeed fail to legitimize many core applications of contemporary antidiscrimination law. To accord with antiperfectionism and the principles of public reason, a conception of harm must appeal to a broad range of conceptions of the good. A conception of harm can achieve this ecumenism, however, only by characterizing harmful behavior at a relatively high level of generality, in terms of features that many forms of seemingly legitimate conduct share. An antiperfectionist version of the harm principle thus faces a dilemma: it can maintain its appeal to a broad range of conceptions of the good by rather indiscriminately subjecting conduct to governmental regulation, thereby compromising its liberalism, or it can distinguish more carefully between legitimate and illegitimate forms of conduct by adverting to other, non-harm-based moral principles, thereby compromising its antiperfectionism. This dilemma is especially acute in the case of discrimination. For all the most plausible candidate accounts of the harm of discrimination—a denial of opportunity, a wrongful setback to one’s interests, a form of psychological injury or stigma, or an affront to one’s dignity—also threaten to condemn and subject to governmental regulation many forms of conduct that liberals regard as legitimate, or at least worthy of toleration. The harm principle fails to deliver on its promise of an antiperfectionist, individualistic account of discrimination.

II. DELIMITING THE HARM PRINCIPLE

Antiperfectionism and the Harm Principle

Before critiquing any doctrine of political legitimacy, we must ensure that our target accurately reflects the convictions of its adherents. The need for such precision is especially pressing in the case of the harm principle, a rather malleable doctrine that political philosophers tend to bend to fit the contours of their broader theories. Indeed, some iterations of the harm principle articulate no doctrine of political legitimacy at all,

but rather identify certain desirable aims of governmental regulation. Consider Joel

Feinberg's initial formulation of the harm principle:

It is always a good reason in support of penal legislation that it would probably be effective in preventing (eliminating, reducing) harm to persons other than the actor (the one prohibited from acting) *and* there is probably no other means that is equally effective at no greater cost to other values.⁴

Instead of specifying the grounds on which the government may legitimately regulate individual behavior, this version of the harm principle merely identifies the prevention of harm to others as a *prima facie* reason to criminalize certain acts. It does not limit the government in its use of political power and thus does not function as a doctrine of political legitimacy. To render Feinberg's initial formulation of the harm principle a doctrine of political legitimacy, we would have to amend it to specify preventing harm to others not simply as a "good reason" to criminalize individual behavior, but rather as the *only* acceptable reason for criminalization (or, at least, one of a finite number of acceptable reasons).⁵ Doctrines of political legitimacy prohibit the government from interfering with individual liberty on certain grounds; they do not proclaim the desirability of various forms of governmental regulation.

Even when construed as a doctrine of political legitimacy, the harm principle can be impressed into the service of myriad general political programs. For example, contemporary political philosophers have employed the harm principle to further both perfectionist and antiperfectionist goals—both to help the government cultivate its citizens' moral virtue and to protect people from such a moralizing political agenda. This chapter will consider an antiperfectionist version of the harm principle. More specifically, the version considered below is antiperfectionist in two respects: in its normative

⁴ Joel Feinberg, *Harm to Others*, vol. 1 of *The Moral Limits of the Criminal Law* (Oxford: Oxford University Press, 1984), 26.

⁵ Feinberg does eventually render his version of the harm principle a (partial) doctrine of political legitimacy precisely in this way, identifying harm to others as one of two legitimate grounds for the criminalization of individual behavior (the other ground being offense to others). See *ibid.*, 14-15.

underpinnings and in its criterion for assessing the legitimacy of governmental regulation. As we shall discover in the next chapter, Joseph Raz rejects any understanding of the harm principle as antiperfectionist in the first respect, instead grounding the doctrine in the value of personal autonomy and deeming that value partially constitutive of the good life.⁶ Like Raz, Feinberg invokes the value of personal autonomy to deprecate legal moralism and to justify the harm principle, but he understands that value not as a constituent of the good life, but rather as a right people enjoy against moralizing interference by the government in their personal affairs, a right whose exercise does not necessarily promote individual well-being.⁷ Indeed, Feinberg attempts to decouple his favored understanding of personal autonomy from any one particular conception of the good life, implicitly defending its conformity with the requirements of public reason when he claims that it “accords uniquely with a self-conception deeply imbedded in the moral attitudes of most people and apparently presupposed in many of our moral idioms.”⁸ Because it has such ecumenical appeal, his understanding of personal autonomy need not presuppose any particular conception of the good and thus need not compromise the antiperfectionism of the broader case for the harm principle.⁹

In addition to forswearing any allegiance to a particular conception of the good life, the version of the harm principle considered in this chapter is antiperfectionist in that it proposes a criterion of political legitimacy—harm to others—that purports to forgo all moral assessment of individual actions and persons’ characters. In particular, an antiperfectionist version of the harm principle insists that the “use [of state coercion] is

⁶ See Joseph Raz, *The Morality of Freedom* (Oxford: Oxford University Press, 1986), ch. 15.

⁷ See Joel Feinberg, *Harm to Self*, vol. 3 of *The Moral Limits of the Criminal Law* (Oxford: Oxford University Press, 1986), 57-62.

⁸ *Ibid.*, 61-62.

⁹ David Richards similarly grounds his version of the harm principle in an antiperfectionist conception of autonomy, one that requires the government to “respect[] the higher-order capacity of the agent to exercise rational autonomy in choosing and revising his ends, *whatever they are*” (David A.J. Richards, *Sex, Drugs, Death, and the Law* (Totowa, N.J.: Rowman & Littlefield, 1982), 9 (emphasis added)).

not conditional on the moral worth of individual actions, but only on the test of discernible harm to other people. Coercive restraint must be independent of considerations of moral worth displayed by the action, or of moral virtue exhibited by the agent.”¹⁰ It identifies harm, moreover, in terms of “values” that “reflect the necessary protection of general goods”¹¹ rather than any thicker conception of human flourishing. Only if behavior “is deleterious from the point of view of a very wide range of conceptions of the good”¹² will an antiperfectionist version of the harm principle pronounce it “harmful.”

If construed as antiperfectionist in both the senses identified above, the harm principle must function as a direct ends-limiting, rather than a means-limiting, doctrine of political legitimacy. For by specifying only the means the government may employ to regulate individual behavior and not the ends it may pursue through such regulation, a means-limiting version of the harm principle leaves the government free to shape its policies according to moral assessments of actions and the persons who perform them. Consider John Stuart Mill’s version of the harm principle, according to which “the only purpose for which power can be rightfully exercised over any member of a civilized community against his will, is to prevent harm to others.”¹³ On initial inspection, this version appears to function as a direct ends-limiting doctrine of political legitimacy, specifying the prevention of harm to others as the sole legitimate end of governmental regulation. But when we consider it in the context of Mill’s broader political theory, we see that it actually licenses the government to pursue a whole range of other ends, including perfectionist ones like promoting citizens’ moral virtue. For example, a government limited by Mill’s version of the harm principle may still use moral suasion to

¹⁰ Wojciech Sadurski, “Joseph Raz on Liberal Neutrality and the Harm Principle,” *Oxford Journal of Legal Studies* 10 (1990): 122.

¹¹ David A.J. Richards, *Toleration and the Constitution* (Oxford: Oxford University Press, 1986), 251.

¹² Brian Barry, *Justice as Impartiality* (Oxford: Oxford University Press, 1996), 87-88.

¹³ Mill, *On Liberty*, 14.

attempt to reform the moral characters of those of its citizens who engage in harmless but otherwise wrongful behavior; it must eschew only certain other means—namely, any form of “power”—in pursuit of perfectionist goals.¹⁴ Such perfectionism, moreover, can infect the operation of the harm principle itself, enabling the government to discriminate among various forms of harmful behavior based on the moral characters of the persons who engage in them. Indeed, a means-limiting version of the harm principle would even permit the government to punish people *because of* the viciousness they evince through their actions—so long as those actions are also harmful. Because antiperfectionism would condemn this punitive rationale, a means-limiting version of the harm principle cannot ensure the doctrine’s antiperfectionism.

Direct ends-limiting versions of the harm principle, by contrast, are perfectly capable of forbidding moralizing forms of governmental regulation and thus can accord fully with antiperfectionism. By explicitly identifying the prevention of harm to others as the legitimate end of governmental regulation, such versions not only set a necessary condition for such regulation, but also restrict the kinds of reasons the government may adduce to support its policies. In particular, they enjoin the government from considering the moral characters of the persons it proposes to punish; the only legitimate consideration is the harm such punishment would forestall. In seeking to prevent harm to others, then, a government constrained by a direct ends-limiting version of the harm principle may not espouse even as its ulterior motive perfectionist aims like the cultivation of its citizens’ moral virtue, aims that go well beyond the antiperfectionist one of protecting its citizens from harm.

The Structure of the Harm Principle

¹⁴ Ibid., 84.

Even if an antiperfectionist version of the harm principle must function as an ends-limiting doctrine of political legitimacy, it might nevertheless attempt to identify harmful conduct—and thus specify its requirements—at a number of different levels of generality. Most straightforwardly, the harm principle can identify harm at the level of discrete acts and require that every instance of a general form of behavior prohibited by a governmental regulation be harmful for that regulation to be legitimate. According to this stringent standard, a regulation will be illegitimate even if only one envisioned instance of the behavior it proscribes is harmless. To see just how demanding this standard actually is, consider the prohibition of breach of contract.¹⁵ The law tends not to distinguish between breaches of contract that harm promisees and those that do not, but rather proscribes all breaches, regardless of their effects on individual promisees. But if we specify the requirements of the harm principle at the level of individual actions rather than at the level of classes of behavior, then the harm principle will preclude regulations as indiscriminate as the law’s approach to breach of contract.

This conclusion prompts some proponents of the harm principle to reformulate the doctrine’s requirements at higher levels of generality, seeking to subject to governmental regulation certain wrongful, but nonetheless harmless, actions. For example, some proponents distinguish between “action-types” and “token actions”—between classes of behavior and their particular instantiations—and then articulate the requirements of the harm principle in terms of the former rather than the latter. So long as there exists a significant probability that a given action-type will be harmful, these proponents maintain, the government may proscribe all its tokens—whether or not each token itself actually causes harm.¹⁶

¹⁵ This example is borrowed from John Gardner and Stephen Shute, “The Wrongness of Rape,” in *Oxford Essays in Jurisprudence: Fourth Series*, ed. Jeremy Horder (Oxford: Oxford University Press, 2000), 196.

¹⁶ For one example of this strategy, see Dudley R. Knowles, “A Reformulation of the Harm Principle,” *Political Theory* 6 (1978): 233-46.

Other proponents of the harm principle abstract from particular acts even further, identifying harm not at the level of token actions or even action-types, but rather in terms of the effects of prohibiting or tolerating a given form of behavior. John Gardner and Stephen Shute, for instance, maintain that the harm principle “forbids the attaching of legal sanctions to wrongdoing except to the extent that this is necessary to prevent harm and proportionate to the harm thereby prevented,” but then consider it “no objection under the harm principle that a harmless action was criminalized, nor even that an action with no tendency to cause harm was criminalized. It is enough to meet the demands of the harm principle that, if the action were not criminalized, *that* would be harmful.”¹⁷ In a similar vein, Colin Bird contends that the harm principle requires only “that (a) serious harms would result from failing to enforce a general prohibition on a class of conduct, and (b) coercively imposing such a prohibition will not itself cause even greater harm.”¹⁸ Bird, like Gardner and Shute, would thus deem harmful for purposes of the harm principle not only behavior that directly causes harm, but also behavior whose remote effects, including those effects stemming from the government’s failure to regulate the behavior, prove deleterious.

By identifying harm in terms of the broader effects of the government’s regulatory decisions, Bird and Gardner and Shute conceive of the harm principle as the kind of two-level doctrine of political legitimacy considered in Chapter 1. That is, while they posit harm prevention as the ultimate criterion of political legitimacy, they do not incorporate this criterion directly into the first-order rules governing the exercise of political power. Rather, they permit the government to prohibit behavior with no direct harmful effects—so long as such prohibitions prevent more harm than they cause overall.

¹⁷ Gardner and Shute, “Rape,” 195, 216.

¹⁸ Bird, “Harm Versus Sovereignty,” 181.

As Chapter 1 argued, antiperfectionist liberalism cannot accommodate these kinds of two-level doctrines of political legitimacy. This inability is especially apparent in the case of versions of the harm principle like Bird's and Gardner and Shute's, whose broad conceptions of harm threaten fundamental liberal commitments in at least two ways. First, if the government can satisfy the demands of the harm principle merely by demonstrating that a prohibition will yield a net reduction of harm, without having also to prove that a given instance of the prohibited conduct is itself harmful, then the government will, in principle, be able to regulate a rather broad range of behavior. To be sure, almost every version of the harm principle, whether one- or two-level, will require that any regulation actually prevent harm, not simply that the government intend or expect it to do so. This requirement, in turn, will effectively insulate from governmental interference much conduct that would otherwise fall within the scope of Bird's and Gardner and Shute's interpretations of the harm principle. Nevertheless, to identify harm in terms of the effects of regulating or tolerating a given form of behavior, rather than in terms of the behavior's more immediate consequences, is to expand the government's writ. This result will disconcert many liberals.

Second, and even more disquieting from an antiperfectionist perspective, this broad conception of harm will subject to regulation behavior that causes harm not necessarily by producing deleterious effects for its instant victims, but by engendering in third parties fear, anxiety, or other unsettled mental states. For example, as construed by Gardner and Shute, the harm principle would license the government to prohibit even those breaches of contract that do not harm individual promisees, since such breaches nevertheless breed distrust and "erode the practice of maintaining reliable voluntary obligations, thereby harming everyone who has reason, or might come to have reason, to

enter into such obligations.”¹⁹ At first glance, one might worry that any conception of the harm principle that so predicates the legitimacy of governmental regulation on people’s moral beliefs threatens individual liberty. Arthur Ripstein, for instance, contends that the harm principle would no longer impose meaningful limits on governmental power if it began “treating people’s *response to* a type of action, or the failure to prohibit it, as a relevant effect on which to base a prohibition.”²⁰ After all, various forms of behavior elicit myriad emotional reactions, reactions that can harm not only the persons who harbor them, but also the shared social practices in which those persons participate. Thus, the argument goes, Gardner and Shute’s version of the harm principle would open a Pandora’s Box of illiberal regulation—especially regulation aimed at curbing so-called private immoralities, which, though perfectly harmless in themselves, can nonetheless engender fear and insecurity and thereby compromise the integrity of important social practices. No version of the harm principle that permits such regulation may rightly proclaim itself liberal.

But while Ripstein correctly highlights the potentially illiberal implications of versions of the harm principle as permissive as Gardner and Shute’s, he fails to consider how we might amend such versions better to safeguard individual liberty. In particular, he seems to assume that we cannot modify the harm principle to discriminate between legitimate and illegitimate emotional responses to objectionable behavior. No intrinsic feature of the harm principle, however, prevents the doctrine from drawing such distinctions. Because the harm principle purports to impose only a necessary, and not a sufficient, condition for legitimate regulation, it in no way compels the government to implement every possible regulation that would yield a net reduction of harm. We thus

¹⁹ Gardner and Shute, “Rape,” 196.

²⁰ Arthur Ripstein, “Legal Moralism and the Harm Principle: A Rejoinder,” *Philosophy & Public Affairs* 35 (2007): 198. For another critique of the harm principle along these lines, see Hamish Stewart, “The Limits of the Harm Principle,” *Criminal Law and Philosophy* 4 (2010): 28-29.

can coherently frame the doctrine to require the government to disregard harms that derive from people's illegitimate biases and prejudices even as it addresses other harms that derive from their legitimate fears and anxieties. *Pace* Ripstein, it seems we need not abandon our commitment to individual liberty in order to bring within the purview of the harm principle at least some harms that stem from people's mental states.

By permitting the harm principle to discriminate between legitimate and illegitimate emotional responses to objectionable behavior, however, we contravene another tenet of contemporary liberalism: its antiperfectionism. For the harm principle cannot make such discriminations without evaluating the morality of both the behavior that elicits the responses and the persons who have them. In particular, to dismiss certain emotional responses as illegitimate and thus harmless, the harm principle must deny that the behavior that arouses those responses is immoral and indict the persons who have them as unjustified. As we have seen, rather than attending solely to the harmful effects of various forms of behavior, a version of the harm principle that so assesses actions and persons adverts to a conception of the good life, contravening the requirements of public reason and disavowing antiperfectionism. Indeed, Gardner and Shute readily concede that their version of the harm principle permits the government to appraise the moral wrongfulness of individual behavior; they merely insist that the government limit its decisions to proscribe certain behavior to considerations of the harm to be prevented rather than more moralistic judgments.²¹ However convincing we might find this attempt to reconcile the harm principle with legal moralism, it purchases greater normative sensitivity at the price of overt moral partisanship, thereby forsaking antiperfectionism for perfectionism.

²¹ Gardner and Shute, "Rape," 196.

If it is to remain antiperfectionist, then, the harm principle cannot bring harmless actions within the government's regulatory purview obliquely by citing the harmful effects stemming from the emotional responses some people might experience were such actions tolerated. An antiperfectionist version of the harm principle cannot, in other words, function as the kind of two-level doctrine of political legitimacy considered in Chapter 1, lest it license the government to make the kinds of moralistic assessments that antiperfectionism forbids.

But we need not foreclose two other "indirect strategies" for justifying the regulation of certain ostensibly harmless actions according to an antiperfectionist version of the harm principle²²—strategies that do not render the doctrine a two-level principle in the sense used in Chapter 1. First, as noted above, even though a given token action might end up being harmless, it might instantiate an action-type that is so likely to cause harm that the harm principle permits the wholesale proscription of all the type's tokens. Similarly, the government will often confront classes of actions whose members appear perfectly harmless when taken individually but that cause serious harm in the aggregate, harm that can be prevented only by regulation that forbids each individual action. Only the most adamant version of the harm principle would condemn such regulation.

Neither of these two indirect strategies renders the harm principle a two-level doctrine of political legitimacy. For each extends the harm principle not by permitting the government to regulate harmless behavior, but rather by refining the concept of harm—through the notions of risk and aggregation, respectively. Each strategy does so, moreover, while preserving the harm principle's antiperfectionism, allowing the government to maintain its focus on harm rather than requiring it to assess the morality of actions or persons. Thus, if it turns out that some instances of discrimination do not harm their

²² For further discussion of these two indirect strategies, see Arthur Ripstein, "Beyond the Harm Principle," *Philosophy & Public Affairs* 34: (2006): 222-24.

immediate victims, we must consider the possibility that an antiperfectionist version of the harm principle can nonetheless subject them to governmental regulation by deploying one of these two indirect strategies.

Whether pursued through a direct or an indirect strategy, the task of providing an account of the harmfulness of discrimination is urgent. For the harm principle permits the government to employ its political power to regulate only that individual behavior that harms others—no matter how wrongful it might be in other respects. If discrimination does not harm its victims, but is instead morally wrongful only for some other reason, then a government limited by the harm principle may not legitimately enact antidiscrimination law.

III. TORT LAW AND ANTIDISCRIMINATION LAW

It seems so natural to suppose that discrimination does indeed harm its victims, and that the harm principle consequently legitimates antidiscrimination law, partly because antidiscrimination law strongly resembles tort law, which is traditionally understood as a measure for preventing harm. As the previous chapter noted, most jurisdictions distinguish between two types of discrimination: “direct” and “indirect” (or “adverse impact”). Recall that, according to U.K. antidiscrimination law, one directly discriminates against another person based on race or sex if, “because of” her race or sex, one treats her “less favourably than [one] treats or would treat others,”²³ a form of discrimination numerous scholars have likened to the “conventional harm-based tort.”²⁴ For example, like the tortfeasor, a person who discriminates directly must act with the requisite mental state and cause his victim’s misfortune if he is to be held legally liable.²⁵ By deeming direct discrimination harmful and therefore placing it squarely within the

²³ Equality Act 2010, § 13(1).

²⁴ John Gardner, “Liberals and Unlawful Discrimination,” *Oxford Journal of Legal Studies* 9 (1989): 4. See also Richard A. Epstein, “The Harm Principle—And How It Grew,” *University of Toronto Law Journal* 45 (1995): 410.

²⁵ Gardner, “Liberals,” 4.

purview of the harm principle, we would confirm its apparent affinity with the harm-based tort.

This affinity, however, seems to belie the preceding construal of the harm principle as an antiperfectionist doctrine of political legitimacy. For, as we just saw, tort law deems alleged wrongdoers liable only after scrutinizing some of the mental states with which they act. Tort law thus seems to contemplate not simply the deleterious effects of individual behavior, but also the morality of that behavior and of the persons who engage in it. If antidiscrimination law shares with tort law this apparent concern with individual moral character, then it seems unlikely we can put antidiscrimination law on a firm antiperfectionist footing merely by adopting an individualistic, harm-based account of the wrongfulness of discrimination.

This conclusion would indeed follow if we could not consider at least some of a person's mental states without adjudging his moral character and miring ourselves in intractable disputes about the nature of the good life. But it is by no means obvious that we must make such judgments and join such disputes to deem a person's wrongful actions intentional, reckless, or negligent—at least as these terms are used in tort law. On the contrary, many philosophers of tort law have argued that we can determine even an actor's *intentions* without having to assess his *motivations*, the reasons for which he acts. Ripstein, for example, contends that determinations of legal liability diverge from those of moral culpability in attending only to “the external aspects of action” rather than to “the quality of a person's will or character.”²⁶ Though denizens of the mind rather than of the external world, one's intentions need not evince the “quality” of one's “will or character,” so long as we understand intentions thinly, to refer to an actor's knowledge or

²⁶ Arthur Ripstein, *Equality, Responsibility, and the Law* (Cambridge: Cambridge University Press, 1999), 4.

awareness of his action's consequences in the external world rather than the reasons on which he acts. So understood, intentions are "external" in Ripstein's sense.

Jules Coleman marks a similar distinction, confirming Ripstein's suggestion that we can subject a person to legal liability and yet balk at a thoroughgoing assessment of his moral character. As Coleman explains, when tort law requires fault at all, it requires "fault in the doing" but not necessarily "fault in the doer"—a failure "to measure up to the relevant standard of conduct" but not necessarily a failure to have cultivated a virtuous character.²⁷ Because "fault in the doing" varies with the degree of knowledge with which the actor acted—in particular, his knowledge of his action's consequences—we cannot impute legal liability to the actor without examining some of his mental states. But such an examination stops well short of a more comprehensive evaluation of the actor's reasons for action and, hence, his moral character. For tort law cares not whether wrongdoing stems from malicious motivations or excused mental states; it is completely indifferent to any "fault in the doer," seeking only to identify the perpetrator of a given wrong through the sheer fact of his decision to act.²⁸ This indifference to persons' moral characters permits tort law to eschew any reliance on a particular conception of the good life and thus to heed the strictures of antiperfectionism—even as it scrutinizes persons' intentions and certain other mental states. Likewise, it seems prohibitions against direct discrimination need not presuppose perfectionism simply because they emulate tort law in espousing a harm-based, intention-sensitive account of wrongdoing.

Can we also understand prohibitions against *indirect* discrimination as manifesting tort law's antiperfectionist concern with preventing and rectifying harm? Initially, indirect discrimination seems to bear little resemblance to the harm-based tort. Again according to U.K. antidiscrimination law, one indirectly discriminates against

²⁷ Jules Coleman, *Risks and Wrongs* (Cambridge: Cambridge University Press, 1992), 220, 217.

²⁸ *Ibid.*, 219.

another person based on race or sex if one “applies” to her a “provision, criterion or practice” that one would apply equally to people of other races or the other sex but that nevertheless disproportionately disadvantages members of her race or sex and for which one can offer no compelling race- or sex-neutral justification.²⁹ Some scholars have attempted to cabin such discrimination into the structure of tort law. Denise Réaume, for example, portrays prohibitions against indirect discrimination as reflecting “an implicit expansion of the concept of fault.”³⁰ Such an “expansion” seems to preserve the analogy between antidiscrimination law and tort law, implicitly employing the tort-law concepts of recklessness and negligence to impose legal liability in cases where the actor *should*, but does not, have knowledge of the consequences of his actions. Both direct and indirect discrimination, then, seem to fit the model of the conventional harm-based tort.

IV. THE HARM OF DISCRIMINATION

Recall that, as construed in this chapter, the harm principle functions as a direct ends-limiting doctrine of political legitimacy, one that permits the government to regulate only harmful conduct. Such an understanding of the harm principle accords with antiperfectionism, as it prohibits the government from attempting to cultivate its citizens’ moral virtue. If antidiscrimination law is to pass muster before the harm principle, so understood, discrimination cannot simply offend some fundamental moral precept; it must also harm its victims, having an appreciably deleterious effect on their well-being. The harm principle thus requires a plausible account of the harm discrimination causes its victims before it can pronounce antidiscrimination law legitimate. Such an account, moreover, must avoid any entanglement with a particular conception of the good life, lest it co-opt the harm principle into the perfectionist project.

Denial of Opportunity

²⁹ Equality Act 2010, § 19.

³⁰ Denise G. Réaume, “Harm and Fault in Discrimination Law: The Transition from Intentional to Adverse Effect Discrimination,” *Theoretical Inquiries in Law* 2 (2001): 351, 353.

We might be tempted to accept as obvious the proposition that discrimination is harmful, since such discrimination necessarily has undesirable material consequences for its victims. Paul Brest, for one, emphasizes such adverse consequences in seeking to vindicate the “antidiscrimination principle,” which proscribes “race-dependent decisions that *disadvantage* the members of minority groups.”³¹ According to Brest, this disadvantage is harmful because it involves “the denial of the opportunity to secure a desired benefit.”³² If such a denial is indeed harmful, rather than merely undesirable or morally objectionable, then Brest will have legitimized antidiscrimination law according to the harm principle.

By deeming the mere denial of an opportunity harmful, however, Brest’s conception of harm licenses the government to encroach rather deeply into its citizens’ personal lives. For, as John Gardner rightly notes, we regularly suffer denials of opportunities; if every denial counted as a harm in the sense relevant for the harm principle, then the government could legitimately regulate conduct that liberals commonly suppose to be a matter of private choice. To fully grasp the illiberal implications of Brest’s conception of the harm of discrimination, consider Gardner’s examples of a person who is denied a job because he is unqualified and a person who is not invited to a dinner party because he is unpopular.³³ In both cases, most liberals would balk at pronouncing these denials of opportunities harmful in the sense relevant for the harm principle, a reluctance that suggests that “the harm principle does not respond to just any disappointment or rejection, but requires something further.”³⁴ If we do not demand “something further,” but instead consider the mere denial of an opportunity to be harmful,

³¹ Paul Brest, “The Supreme Court 1975 Term—Foreword: In Defense of the Antidiscrimination Principle,” *Harvard Law Review* 90 (1976): 2 (emphasis added).

³² *Ibid.*, 8.

³³ Gardner, “Liberals,” 6.

³⁴ *Ibid.*

then the harm principle will not function as the limit on political power that liberals have traditionally supposed it to be.

We can glean from Gardner's examples two distinctions the harm principle must draw if it is effectively to limit political power. First, as the job-application example suggests, the harm principle must distinguish between legitimate and illegitimate denials of opportunities, between denials that simply disadvantage their victims and those that wrong them as well. To be sure, the harm principle need not rely exclusively on its own conceptual resources to draw such a distinction, but can enlist the aid of independent moral principles. But if these other moral principles are not to compromise the harm principle's antiperfectionism, they must share the harm principle's singular focus on the victims of discrimination rather than the perpetrators. Without adverting to perfectionist justifications, then, the harm principle and any ancillary moral principles together must explain why the government may prohibit one denial of an opportunity but not another—even when both denials have identical material consequences for their respective victims.³⁵

Second, the harm principle must distinguish between more and less significant denials of opportunities, for even some wrongful denials seem to be inappropriate objects of governmental regulation. In another version of Gardner's dinner party example, for instance, the host might commit a wrong by omitting someone from his guest list for an indisputably illegitimate reason, but we nonetheless tend to think that the government must abide such wrongdoing. This intuition seems to derive much of its force from the relative triviality of the denied opportunity: few believe that the government may abridge individual liberty to secure for some of its citizens an opportunity, like that of attending a

³⁵ Perhaps Brest means to distinguish between legitimate and illegitimate denials of opportunities with the notion of "race-dependent decisions." But of course, Brest would then have to provide an independent account of why such decisions are illegitimate, an account that maintains the harm principle's antiperfectionism. This chapter considers the possibility of such an account in the context of Joel Feinberg's interpretation of the harm principle.

dinner party, whose denial usually detracts little from one's well-being. To be sure, we also worry about the context in which the denial occurs—namely, an intimate relationship within the home. But intimacy poses no absolute bar to governmental interference; when the denial of opportunity is severe enough—when, to take an extreme example, one person threatens the physical safety of a friend or other intimate relation within the home—we do not hesitate to authorize the government to intervene. The harm principle must therefore distinguish not only between legitimate and illegitimate denials of opportunities, but also between trivial and more significant forms of disadvantage. By defining the harm of discrimination as the mere denial of an opportunity, regardless of the denial's legitimacy or significance, Brest's "antidiscrimination principle" neglects both distinctions and, in so doing, compromises the liberal credentials of the harm principle.³⁶

Wrongful Setback to Interests

To maintain the harm principle as a meaningful constraint on political power, we must develop a more refined account of the harm of discrimination, an account that distinguishes the pernicious effects of discrimination from other forms of misfortune that no liberal government may legitimately attempt to rectify. Joel Feinberg's moralized conception of harm as a "wrongfully set-back interest"³⁷ seems to provide just such an account.

To understand Feinberg's conception of harm, we must explicate each of its elements. According to Feinberg, a person has an interest in something if he "has a stake" in it, if he "stands to gain or lose depending on [its] nature or condition."³⁸ Interests thus constitute "distinguishable components of a person's well-being: he flourishes or

³⁶ For an argument that the harm principle lacks the conceptual resources to draw these distinctions, see Bernard E. Harcourt, "The Collapse of the Harm Principle," *Journal of Criminal Law & Criminology* 90 (1999): 109-94.

³⁷ Feinberg, *Harm to Others*, 105 (emphasis omitted).

³⁸ *Ibid.*, 33-34.

languishes as they flourish or languish.”³⁹ But some interests are more integral “components of a person’s well-being” than others. In particular, Feinberg identifies three categories of increasingly important, or “ulterior,” goals a person might have, which, in turn, correspond to three categories of increasingly important interests. Our least ulterior interests are those we have in satisfying our “*instrumental wants*,” our “immediate wants whose fulfillments . . . are linked up, either as means or as necessary conditions, to the advancement of more ulterior goals.”⁴⁰ The second class of interests comprises our “*welfare interests*,” whose joint satisfaction guarantees “that minimum level of physical and mental health, material resources, economic assets, and political liberty that is necessary if we are to have any chance at all of achieving our higher good or well-being, as determined by our more ulterior goals.”⁴¹ Finally, our most important interests are our “ulterior interests” in realizing our “*focal aims*,” or dominant ends, which we value not only instrumentally as means to still further ends, but also intrinsically as ends in themselves.⁴²

A person is harmed, Feinberg maintains, only if one of his interests is “set back.” Acknowledging that the notion of a setback is somewhat vague, Feinberg attempts to elucidate it through a “kind of directional metaphor,” suggesting that to set back an interest “is to reverse its course, turn it away, put it back toward the point from which it started. In terms of its associated goals, it is to reverse its progress, to put it in a worse condition than it was formerly in.”⁴³ But we should not take these metaphors too literally, for, as Feinberg explains, one person, *A*, can thwart the interest of another person, *B*, not only when *A* makes *B* worse off than *B* was, but also when *A* makes *B* worse off than *B*

³⁹ Ibid., 34.

⁴⁰ Ibid., 56-57 (emphasis omitted).

⁴¹ Ibid., 57.

⁴² Ibid., 60.

⁴³ Ibid., 53.

would have been had *A* not intervened.⁴⁴ Feinberg's definition of harm thus incorporates a "counterfactual test," according to which *A* sets back *B*'s interest if "*B*'s personal interest is in a worse condition . . . than it would be in had *A* not acted as he did."⁴⁵ Indeed, on this view, one can set back another person's interest even by improving it—just not by as much as it would have been improved had one not intervened.⁴⁶

While the notion of a setback, so defined, is "relativistic," implying the deterioration of an interest from a given "starting point" or "baseline" (even if a counterfactual one), it also has an absolute dimension, for we can distinguish particularly acute setbacks from less significant ones.⁴⁷ As we discovered when considering Brest's "antidiscrimination principle," the harm principle must draw such a distinction if it is to sufficiently limit governmental interference with individual liberty. Feinberg eventually makes clear that the harm principle cognizes not every wrongful setback to an interest, but only those wrongful setbacks that exceed a certain threshold of severity. A setback to a welfare interest, for example, is not harmful in the sense relevant for the harm principle unless it depreciates the interest "below a tolerable minimum level."⁴⁸ More generally, according to one of the "mediating maxims" with which Feinberg tempers the harm principle's legitimating force, the government may not seek to prevent "bare minimal invasions of interest just above the threshold of harm," lest, in so doing, it cause even greater harm to wrongdoers' legitimate interests—especially their interest in liberty.⁴⁹

As the failure of Brest's account of the harm of discrimination also demonstrated, the harm principle must distinguish not only between trivial and significant denials of opportunities, but also between legitimate and illegitimate ones. Feinberg appreciates this

⁴⁴ See Joel Feinberg, *Freedom and Fulfillment* (Princeton: Princeton University Press, 1992), 7.

⁴⁵ *Ibid.*

⁴⁶ *Ibid.*, 7-8.

⁴⁷ Feinberg, *Harm to Others*, 54.

⁴⁸ *Ibid.*, 206.

⁴⁹ *Ibid.*, 188, 190.

imperative and consequently includes among the harms cognizable by the harm principle only those significant setbacks to interests that are also “wrongful,” where

[o]ne person, *A*, can be said to *wrong* another, *B*, when he treats him unjustly. More precisely the injustice occurs when *A*’s act or omission has as its intention to produce an adverse effect on *B*’s interest, or is negligent or reckless in respect to the risk of such an effect; and *A*’s conduct is morally indefensible [i.e., neither justifiable nor excusable]; and *B*’s set-back interest is one that he has a *right* to have respected.⁵⁰

A person has a right, moreover, when one of his interests grounds a “valid claim” either against particular persons to their performance of particular actions or against all persons to “noninterference in his private affairs.”⁵¹ Therefore, on Feinberg’s account, one person harms another in the sense relevant for the harm principle when he sets back one of the other person’s interests and, in so doing, also violates one of the other person’s rights. Adopting this moralized conception of harm, Feinberg’s version of the harm principle licenses the government to prevent not all setbacks to interests, but only “that subclass of harms that are wrongs.”⁵²

We might begin to worry that, in so predicating the legitimacy of governmental regulation on the morality of the behavior to be regulated, Feinberg’s version of the harm principle abandons all pretensions to antiperfectionism, adverting to a conception of the good life. Any discomfort seems only to intensify once we realize that Feinberg’s conception of harm embodies a notion of wronging that requires us to scrutinize persons’ mental states. Like the notion of wronging presupposed by much of tort law, Feinberg’s incorporates a “*mens rea* principle,” according to which the wrongfulness of an action depends partly on the degree of knowledge with which the actor performs it.⁵³ As we have seen, this principle requires no assessment of persons’ moral characters and hence

⁵⁰ Ibid., 107-08.

⁵¹ Ibid., 109.

⁵² Ibid., 110.

⁵³ John Gardner, “Wrongs and Faults,” in *Appraising Strict Liability*, ed. A.P. Simester (Oxford: Oxford University Press, 2005), 69.

no commitment to a particular conception of the good life. But unlike tort law's notion of wrongdoing, Feinberg's comprehends only "fault-anticipating wrongs," "wrongs that are committed only in the absence of justification or excuse—in other words, only in the presence of fault."⁵⁴ Adopting this focus conjures the specter of perfectionism, for it seems that we cannot determine whether someone acts with justification or excuse unless we assess the reasons for which he acts, which, in turn, requires us to appraise his character, those personal motivations and other traits that might exculpate him for his alleged wrongdoing. Apparently breaching the distinction between "fault in the doing" and "fault in the doer," Feinberg's notion of wrongdoing seems to compromise the harm principle's antiperfectionism.

Indeed, in his later work, Feinberg explicitly confirms this suspicion, conceding that regulations sanctioned by his version of the harm principle endeavor not only to prevent harm, but also publicly to censure persons who harm others.⁵⁵ This embrace of perfectionism explains Feinberg's focus on "fault-anticipating wrongs": if Feinberg's version of the harm principle authorizes the government to censure certain wrongdoers, to communicate moral blame, then it may recognize as wrongs only those actions for which persons can be morally blameworthy—that is, actions performed with neither justification nor excuse.⁵⁶ Despite its grounding in an ostensibly antiperfectionist conception of autonomy, Feinberg's version of the harm principle nonetheless ends up consorting with perfectionism.

It thus appears that Section II, above, incorrectly depicted Feinberg's version of the harm principle as an antiperfectionist doctrine of political legitimacy. But it would be precipitous to conclude that we may no longer consider Feinberg's conception of harm as

⁵⁴ Ibid., 67.

⁵⁵ Joel Feinberg, *Harmless Wrongdoing*, vol. 4 of *The Moral Limits of the Criminal Law* (Oxford: Oxford University Press, 1990), 12-13.

⁵⁶ Cf. Feinberg, *Harm to Others*, 108-09.

a candidate antiperfectionist account of the harm of discrimination. For the features that render Feinberg's version of the harm principle perfectionist are particularly salient in the context of the criminal law, the context to which Feinberg explicitly confines the doctrine's application.⁵⁷ More specifically, Feinberg limits his account of harm to fault-anticipating wrongs because he endows the criminal law with the so-called expressivist function of censuring wrongdoers and communicating blame. By contrast, as noted above in Section III, in supposing discrimination to be harmful and consequently invoking the harm principle to legitimate antidiscrimination law, theorists assimilate the legislation not to criminal law, but to tort law, a branch of law that arguably performs no expressivist function and hence largely ignores moral culpability. If antidiscrimination law is truly an analogue of tort law rather than criminal law, then it matters not whether those who discriminate act with either justification or excuse; their wrongdoing suffices to subject them to legal liability, irrespective of their moral culpability. Given the apparent affinity between tort law and antidiscrimination law, we should consider the possibility that, once purged of its perfectionist concern with justifications and excuses, Feinberg's conception of harm can ground a plausible account of the harm of discrimination while heeding the strictures of antiperfectionism.⁵⁸

Once it jettisons the concepts of justification and excuse, how is Feinberg's conception of harm to distinguish between legitimate and illegitimate denials of opportunities, as it must do if it is to avoid having the illiberal implications of Brest's "antidiscrimination principle"? The most obvious basis for distinction is another element of Feinberg's account of harms: the requirement that a setback to a person's interest also violate one of the person's rights. It is unclear, however, whether the concept of rights

⁵⁷ See *ibid.*, 3-4.

⁵⁸ Indeed, Feinberg frequently does the purging for us, automatically deeming behavior that sets back interests to be harmful without considering, as his own account of harm requires, whether the behavior is also wrongful. See Hamish Stewart, "Harms, Wrongs, and Set-Backs in Feinberg's *Moral Limits of the Criminal Law*," *Buffalo Criminal Law Review* 5 (2001): 47-67.

preserves any meaningful role for the harm principle in the liberal account of political legitimacy or whether it instead effaces the doctrine altogether. We can understand Feinberg's account of rights in one of two ways. First, Feinberg's account of rights might supervene on his account of harm, so that a person's rights are violated whenever his interests are set back. On this view, rights are not independent sources of moral entitlement, but "normative advantages that we give people for the purpose of protecting their interests."⁵⁹ The problem with this understanding of rights is that it is superfluous: if a setback to a person's interests necessarily violates his rights, then the concept of rights in no way restricts the class of setbacks that count as harms for purposes of the harm principle; all setbacks continue to count as harms, and the harm principle functions as if the concept of rights did not exist.⁶⁰ Though perhaps a useful shorthand for denoting setbacks to interests, this understanding of rights fails to cabin the harm principle's legitimating force.

According to a second understanding of Feinberg's account of rights, rights are defined independently of setbacks to interests, so that a setback need not entail a rights violation. More specifically, this understanding posits a category of "Kantian" or "juridical rights that a person has just in virtue of being a person who is subject to the law."⁶¹ Because these Kantian rights are grounded in our personhood rather than our interests, they can limit the application of the harm principle by distinguishing among setbacks to interests. Thus, when supplemented by a doctrine of Kantian rights, the harm principle authorizes the government to regulate not all setbacks to interests, but only certain ones—namely, those setbacks that also violate Kantian rights.

We might wonder, however, about the character of these Kantian rights. In particular, are they perfectionist or antiperfectionist? Do they presuppose a particular

⁵⁹ Stewart, "Limits," 19.

⁶⁰ See *ibid.*, 23.

⁶¹ *Ibid.*, 19.

conception of the good life, or can they be explicated solely in the terms of public reason? Whether perfectionist or antiperfectionist, Kantian rights fail to preserve the harm principle as a meaningful limit on political power. Obviously, if Kantian rights are perfectionist, then we may not invoke them to supplement an antiperfectionist version of the harm principle, whose primary purpose is to extricate questions about the legitimate scope of governmental power from questions about the good life. If Kantian rights are instead antiperfectionist, then they do not merely supplement the harm principle, but supplant it altogether. For once we recognize a category of antiperfectionist Kantian rights, it is difficult to explain why we should not authorize an antiperfectionist government to remedy violations of those rights, regardless of whether the rights violations also cause harm, thus leaving the harm principle with no independent liberty-protecting role to play.⁶² To be sure, the harm principle might continue to identify harmful conduct as an important object of governmental concern. But if it uses the notion of Kantian rights to distinguish between legitimate and illegitimate denials of opportunities, the harm principle will authorize the government to remedy only those harms that also violate Kantian rights. Consequently, the doctrine of Kantian rights, not the harm principle, will limit the exercise of governmental power; the doctrine of Kantian rights, not the harm principle, will function as a doctrine of political legitimacy.⁶³

The failure of Feinberg's account of rights and wrongdoing to meaningfully limit his version of the harm principle (without supplanting it altogether) places even more weight on another component of his theory that we have already noted: the "mediating maxim" according to which only significant setbacks to interests count as harms for purposes of the harm principle. Indeed, that mediating maxim can be understood to distinguish indirectly (and imperfectly) between legitimate and illegitimate setbacks to

⁶² Cf. *ibid.*, 18.

⁶³ Chapter 6 will construct a rights-based account of the wrong of discrimination and consider whether such an account can remain antiperfectionist.

interests, the very function that Feinberg's account of rights purports to perform directly. To be sure, as understood by Feinberg, the mediating maxim only determines the magnitude of setbacks to interests so that the setbacks can be balanced against the infringement of liberty that any governmental regulation of individual conduct necessarily entails. But a setback's magnitude and legitimacy are not wholly independent variables. Rather, we tend to view more significant setbacks as more illegitimate, for while it may very well be unfair for one person to impose modest costs on another through his conduct, it is even more unfair when the imposed costs are great. The distinction between significant and modest setbacks thus serves as a rough proxy for the distinction between illegitimate and legitimate (or less illegitimate) ones. This proxy, moreover, can accord with antiperfectionism, for conceptions of the good will all tend to pronounce the same significant setbacks to be harmful, even as they disagree in borderline cases involving more modest setbacks. For example, all conceptions of the good will agree that murder harms its victims, even if they disagree about whether some injury that forecloses only one particular option is illegitimate and thus harmful. By focusing on significant harms alone, then, the harm principle can not only limit governmental interference with individual liberty, but also avert a significant amount of moral disagreement and thereby hew more closely to antiperfectionism.

Within Feinberg's framework, an account of the harm of discrimination must begin by identifying the interests set back by the conduct, those aspects of the victims' well-being that are impaired by discriminatory acts. If Feinberg's version of the harm principle is to legitimate antidiscrimination law, then discrimination must set back its victims' "welfare interests," the only interests that can be infringed directly. By contrast, one person can infringe another's "ulterior interests" only indirectly, by infringing the welfare interests on which the ulterior interests depend—that is, by infringing those basic

interests that one must satisfy to some minimum degree if one is to advance one's further interests.⁶⁴

In fact, the interests Feinberg includes among our welfare interests seem to be precisely the kinds of interests that discrimination infringes. Our welfare interests are, as noted above, those interests we have "in achieving and maintaining that minimum level of physical and mental health, material resources, economic assets, and political liberty that is necessary if we are to have any chance at all of achieving our higher good or well-being."⁶⁵ Discrimination in employment and the provision of public accommodations appears directly to imperil such interests, denying its victims access to the sources of income and material goods we all require to satisfy even our most basic needs. Perhaps even more starkly, because discrimination necessarily forecloses options for its victims, it also seems to impede their welfare interest in liberty, their "interest in having as many *open options* as possible with respect to various kinds of action, omission, and possession."⁶⁶ Antidiscrimination law promises to safeguard this interest, as well as the other welfare interests, against the discriminatory conduct of employers and purveyors of public accommodations.

But does antidiscrimination law restrict its attention to those setbacks that are sufficiently deleterious both to warrant the considerable costs to liberty that attend governmental regulation of individual conduct and to elicit the condemnation of a broad range of conceptions of the good? The significance of a setback to an interest depends on a number of factors, including "the extent of the adverse effect" it has on the interest, the "duration" of that adverse effect, and "the extent of the period during which the interest is in a worse condition than it otherwise would be in."⁶⁷ Each of these factors, in turn, varies

⁶⁴ Feinberg, *Harm to Others*, 62, 112.

⁶⁵ *Ibid.*, 57.

⁶⁶ *Ibid.*, 207.

⁶⁷ Feinberg, *Freedom and Fulfillment*, 9-10.

with the context in which the setback occurs. For example, the fewer options a person continues to enjoy after suffering a setback to one of his interests, the greater the extent of the setback's adverse effect on that interest. If a setback forecloses only one of many more-or-less comparable options for realizing a particular interest, then the setback does not have a significant adverse effect. If, by contrast, a setback to an interest forecloses many options (or one member of a very limited set of options), so that it will be much more difficult for the victim to realize the interest than it would have been without the setback, then the setback does have a significant adverse effect. In short, the significance of a setback varies with the number and quality of the other options the setback eliminates or leaves intact.

We can thus determine whether particular instances of discrimination cause significant, rather than merely *de minimis*, harm only by attending to the broader contexts in which they occur, only, that is, by considering the full range of opportunities victims of discrimination face, both before and after their mistreatment. Indeed, it seems that, on Feinberg's conception of harm, the very same discriminatory act can be either harmful or harmless in the sense relevant for the harm principle, depending on its victim's other circumstances. To take an extreme example, consider a society that broadly resembles that of the southern United States during the era of Jim Crow, the *de jure* regime of segregation and discrimination that relegated blacks to a vastly inferior social status for nearly a century. Because this kind of systematic, state-sanctioned discriminatory apparatus pervades nearly every aspect of social and economic relations, enshrining a caste system that frustrates nearly all social and economic mobility, it deprives its intended victims of the means needed to secure a minimally decent life. In such conditions of privation, almost every instance of discrimination—whether direct or indirect—will harm its victim, significantly setting back one or more of his welfare

interests. For where else can he turn to acquire the resources discriminatory employers and purveyors of public accommodations deny him? How can he offset the setback to his welfare interests wrought by the discrimination? If a person already faces a meager range of opportunities for satisfying his welfare interests, then any act that forecloses yet further opportunities necessarily sets back his welfare interests significantly and thus harms him.

To be sure, some employers and purveyors of public accommodations will flout the prevailing racist mores of their society—a modicum of nonconformity that might mitigate the misfortune of some victims of discrimination, thereby rendering some discriminatory acts less harmful or even harmless. But Feinberg’s version of the harm principle (like any plausible version of the doctrine) can easily countenance the proscription of even these anomalous harmless instances of discrimination by employing the indirect types-tokens strategy identified above in Section II. More specifically, as Feinberg himself notes, the harm principle does not require that every potential violation of a given regulation cause harm for that regulation to be legitimate, but instead permits the government to rely on “empirical generalizations about the likely effects on protected standard interests of various standard kinds of threatening actions” in deciding which actions to prohibit.⁶⁸ For example, the harm principle allows the government to prohibit all instances of drunk driving, even though only some of them turn out to be harmful. Because all acts of drunk driving carry a significant *risk* of causing harm, the harm principle pronounces each harmful in the sense relevant for the harm principle, irrespective of whether the feared harm materializes in a given instance. Similarly, the harm principle will legitimize antidiscrimination law so long as discriminatory acts carry a significant risk of harming their victims, a condition obviously satisfied in an apartheid society like that of the southern United States under Jim Crow. At least in such a

⁶⁸ Feinberg, *Harm to Others*, 190.

pervasively racist society, discrimination is “dangerous by nature,”⁶⁹ especially likely to depress its victims’ welfare interests below a minimum threshold of well-being.

Feinberg’s version of the harm principle authorizes the government to prevent harm by prohibiting such risky behavior.

But discrimination is not so likely to harm its victims in a society that has largely eradicated racial oppression. Although contemporary western liberal democracies are a long way from reaching this state of affairs, we can envision a hypothetical society that has ameliorated its moral culture to such an extent that most discriminatory acts perpetrated within it do not set back their victims’ interests significantly and thus are no longer harmful in the sense relevant for Feinberg’s harm principle. If a substantial majority of employers and purveyors of public accommodations in such a society eschew discrimination, victims of discrimination will still enjoy alternatives to the opportunities withheld by incorrigible racists and sexists. A female job applicant, for instance, will be able to turn to nondiscriminatory firms after being rejected by an anomalous discriminatory one. Thus, while particular instances of discrimination in such a society will continue to set back their victims’ welfare interests, they will usually be insufficiently deleterious to satisfy Feinberg’s version of the harm principle and thus to warrant antidiscrimination law’s interference with individual liberty. Nor will discrimination as an action-type be inherently dangerous—so likely to cause harm that the harm principle permits the proscription of all its various instantiations, harmful and harmless alike. In a society whose citizens largely renounce discrimination, Feinberg’s version of the harm principle appears to deem the core of antidiscrimination law not merely superfluous, but illegitimate.

⁶⁹ Ripstein, “Beyond the Harm Principle,” 223.

Feinberg's version of the harm principle will likely render a similar, though more nuanced, verdict in a society where only a subset of racial minorities or women face pervasive, as opposed to merely sporadic, discrimination. Consider a society that has made substantial but uneven progress in combating racism and sexism. In such a society, equality's benefits will be distributed unevenly, vastly enhancing the life prospects of some members of previously oppressed groups while bypassing others altogether. For example, despite having eradicated de jure segregation and the other formal trappings of apartheid, such a society might tolerate the de facto segregation of significant minority populations within urban ghettos, bastions of poverty and social blight that entrap their inhabitants in cycles of disadvantage and despair.⁷⁰ The denizens of such enclaves will tend to confront material and social conditions akin to those faced by the victims of Jim Crow, conditions woefully inadequate to leading a minimally decent life. Like the discriminatory acts perpetrated in the southern United States under Jim Crow, those perpetrated in these urban ghettos will be sufficiently harmful to warrant governmental regulation according to Feinberg's version of the harm principle.

But suppose that not all racial minorities are consigned to these redoubts of privation and prejudice. Suppose instead that the minority population assumes a bipolar distribution, with one subpopulation residing in urban ghettos and another residing in communities that offer a wide range of opportunities and that tolerate only the most tepid expressions of racism (if they tolerate any such expressions at all), thereby insulating their inhabitants from the more deleterious effects of discrimination. While we would rightly decry discrimination in the latter communities as readily as we would in the former, we cannot automatically conclude that discrimination will harm its victims in both settings. For, as we have seen, Feinberg's version of the harm principle cognizes only those

⁷⁰ See Tommie Shelby, "Justice, Deviance, and the Dark Ghetto," *Philosophy & Public Affairs* 35 (2007): 126-60.

setbacks that depreciate persons' welfare interests significantly, a condition unlikely to be satisfied outside the urban ghettos of the hypothetical society under consideration. In such a society, then, Feinberg's version of the harm principle will justify not the uniform, continuous antidiscrimination regimes with which we are familiar, but rather, at most, only a more fragmentary framework that extends its protections in some communities while withholding them in others.

It might be objected that the preceding criticism considers only the discrete effects of individual discriminatory acts, taken in isolation from one another, while ignoring the cumulative effects that many such acts might have in the aggregate. That is, even if a single act of discrimination does not harm its immediate victim, its effects might combine with those of other, independent discriminatory acts to distort economic markets, fostering general socioeconomic conditions that harm other members of the victim's racial group or sex. Such distortion might occur through something like the following mechanism. If enough employers refuse to hire, say, black employees, then black people will necessarily enjoy fewer employment opportunities and will consequently face keener competition for those opportunities that remain open to them. Although the victim of a particular act of discrimination might ultimately obtain a comparable employment opportunity with a nondiscriminatory employer, and thus might not be harmed in the sense relevant for the harm principle, that act of discrimination might help to foster a labor market in which more black jobseekers are unemployed than otherwise would be. Thus, taken in the aggregate, discrete acts of discrimination that seem harmless in themselves can nevertheless seriously harm racial minorities and women.

This possibility shows how discrimination might constitute an "accumulative harm," a harm that stems not from the isolated effects of any single instance of a

particular kind of conduct, but rather from the aggregate effects of many such instances.⁷¹ In this respect, discrimination might resemble many kinds of environmental pollution. Just as a single factory's toxic discharge will often have no discernible deleterious impact on any of a lake's wildlife, so a single employer's discriminatory hiring practices will often leave all the racial minorities and women in the relevant labor market, including the discrimination's instant victims, unscathed. But if enough factories release toxic chemicals into a lake, then the concentration of the chemicals in the lake might reach harmful levels. Similarly, if enough employers discriminate against black jobseekers, then the labor market might become sufficiently distorted to impede black people's job prospects. In each case, the resulting harm is the product of the concurrent (if uncoordinated) acts of multiple actors, rather than the discrete acts of any single one.

Such aggregate harm can, in turn, have even broader effects, well beyond its immediate impact. For example, even if a particular pollutant directly harms only one species of fish, that species' well-being might be inextricably intertwined with the well-being of other wildlife, so that any harm to the former indirectly injures the latter. Similarly, widespread discrimination against even just the most economically disadvantaged members of a racial group might diminish the social standing (if not the material well-being) of all the group's members by helping to perpetuate unfavorable race-based stereotypes. The upshot is that we cannot appreciate the full extent of the harm of discrimination—and thus the full force of antiperfectionist justifications for antidiscrimination law—without attending to the aggregate effects of many discriminatory acts.

Once we recognize discrimination as an accumulative harm, the antiperfectionist case for antidiscrimination law seems to be much more robust than we initially supposed.

⁷¹ For more on “accumulative harms,” see Feinberg, *Harm to Others*, 225-32.

For it no longer matters that discrete discriminatory acts prove harmless for their instant victims; so long as those acts produce effects that combine to distort economic markets, the harm principle can readily condemn all instances of discrimination and justify the full extent of antidiscrimination law without having to advert to perfectionist notions such as moral character. Indeed, like the systemic account considered in the previous chapter, an accumulative account of the harm of discrimination seems especially congenial to antiperfectionism, abstracting from individual instances of discrimination—and thus from moralistic features of their perpetrators—and focusing instead on the broader impact many such instances can have on society. It is thus too quick to infer the absence of an antiperfectionist, harm-based justification for antidiscrimination law from the fact that certain instances of discrimination fail to harm their victims; only after considering the accumulative effects of all instances of discrimination can we draw such a conclusion.

Although an accumulative account of the harm of discrimination apparently justifies a more extensive antidiscrimination regime than does an account that focuses solely on discrete discriminatory acts, the former account still has limits. Most significant, its validity depends on a rather speculative empirical assumption—to wit, that, in the absence of antidiscrimination law, discrimination will proliferate widely enough for its effects to distort the broader economy. Only on this assumption can the harm principle recognize accumulative harms while still meaningfully circumscribing governmental power, rather than affording the government free rein to regulate individual conduct. Thus, before the government may restrict individual liberty in the name of forestalling an accumulative harm, the harm principle requires it to demonstrate a relatively high

probability that, in the absence of regulation, enough people will engage in the relevant conduct for the accumulative harm actually to materialize.⁷²

Were the harm principle instead to countenance regulation based on a more speculative prediction, two illiberal results would ensue. First, the government would be able to regulate a broad swath of individual conduct virtually on its own say-so, so long as it could cite *some* possibility of harm, however remote, in the absence of regulation. The harm principle would thus cease to impose meaningful limits on governmental power, as liberals suppose it to do. Second, and more troubling from an antiperfectionist perspective, the harm principle would allow views about the moral wrongfulness of individual conduct to infect determinations of the legitimate scope of governmental power. For when the probability of a given harm's materializing is fairly low, the decision whether to regulate conduct that, should it become sufficiently pervasive, threatens to produce the harm ultimately turns on a judgment about the conduct's morality, the kind of judgment that antiperfectionism excludes from the political domain. Some people would deem the probability of harm too remote to warrant regulating the relevant conduct, while others would reach the exact opposite conclusion. The government could not adjudicate between these competing positions without taking a stand on the value of the relevant conduct, without deciding, in other words, whether the conduct was sufficiently valuable to warrant tolerating some risk of harm. Antiperfectionism forbids the government to make such judgments. By contrast, by limiting its attention to more probable harms, the harm principle authorizes only regulation that all conceptions of the good agree are warranted, regardless of their general views of the value of the regulated conduct. The harm principle can thus avert the two illiberal results canvassed in this paragraph only by

⁷² Cf. *ibid.*, 226 ("If, as a matter of fact, not enough persons will act in the way in question to affect adversely any public interest, then there can be no harm in permitting those actions, even if their general or widespread performance would be disastrous.").

requiring that conduct be rather likely to become sufficiently pervasive to cause accumulative harm before authorizing the government to regulate it.

In the hypothetical society we have been considering, discrimination, conceived of as an accumulative harm, does not satisfy this criterion. For discrimination's effects will not accrete to harmful levels in the markets for labor and for goods and services unless a critical mass of employers and purveyors of public accommodations discriminate. If, by contrast, a society can confine discrimination to a few recalcitrant racists and sexists, then the various mechanisms of the economy will likely remain more or less unblemished by any discriminatory taint. In such a society, the prospect of any accumulative harm stemming from discrimination would be too remote to be recognized by the harm principle.

This argument, like the others preceding it, retains the various conditions Feinberg imposes on his version of the harm principle in order to temper the doctrine's legitimating force, to prevent the government from prohibiting conduct that, though technically harmful, is too inconsequential to warrant interference with individual liberty. And yet, someone broadly sympathetic to Feinberg's project might attempt to bolster the harm principle by relaxing one of those conditions—namely, the “mediating maxim” according to which the government may regulate only behavior that causes *significant* harm. Because all discrimination so flagrantly wrongs its victims, we might think the harm principle should countenance the proscription of all discriminatory acts—even in contexts where such acts are likely to cause only trivial harm. Were the harm principle to disregard these more innocuous instances of discrimination, the argument goes, it would permit many racists and sexists to indulge their prejudices with near impunity, to discriminate against racial minorities and women so long as their victims were sufficiently well-off to weather exclusion from a particular employment opportunity or public accommodation.

Why, we might wonder, should antidiscrimination law protect only relatively disadvantaged racial minorities and women, leaving more fortunate ones to fend for themselves against discriminatory employers and purveyors of public accommodations? Surely, the harm principle cannot require such a bifurcated antidiscrimination regime.

By endorsing this line of argument, however, we would transmute the harm principle from a regulation-restricting doctrine of political legitimacy into an expansive mandate for the government to prohibit myriad forms of morally objectionable behavior. For we frequently suffer relatively insignificant wrongful setbacks to our interests, and construing the harm principle to license the government to remedy each of them would eviscerate the doctrine's purported protections of individual liberty. To permit the government, for example, to proscribe "repeated rude and disrespectful remarks to parents, spouses, teachers, and others who have a right to better treatment" simply because "such discourtesies not only 'wound feelings,' but indirectly harm the interest in personal efficiency by causing depression and anger sufficiently great to distract and debilitate," would be to grant the government much broader latitude to regulate individual conduct than liberals suppose it to have.⁷³ Unless we follow Feinberg in qualifying the harm principle with various "mediating maxims," the harm principle will fail meaningfully to limit political power, leaving the government free to encroach deeply into its citizens' private lives.

The significant-harms mediating maxim not only minimizes governmental interference with individual liberty, but also preserves the antiperfectionism of the rationale for such interference. As we have seen, the harm principle's proponents have traditionally conceived of the doctrine as a means of extricating determinations of the legitimate scope of governmental power from disagreements about the morality of

⁷³ Ibid., 188-89.

individual conduct and, more generally, the nature of the good life. The concept of harm, it is thought, is sufficiently ecumenical to support broad agreement among conceptions of the good about which individual conduct merits governmental regulation. But such agreement becomes more elusive as the concept of harm expands to encompass increasingly modest setbacks to interests, which different conceptions of the good will condemn with different degrees of urgency. For Feinberg's version of the harm principle to jettison its significant-harms mediating maxim would thus be for it to abandon its antiperfectionism, requiring the government to side with some conceptions of the good over others on the question of which modest setbacks to prevent rather than simply to ratify the broad agreement among conceptions of the good regarding the permissibility of preventing significant setbacks. Antiperfectionism precludes such moral partisanship. Feinberg's version of the harm principle must continually check its own legitimating capacity if it is to maintain its liberal and antiperfectionist bona fides, a degree of restraint that effectively condemns as illegitimate many applications of antidiscrimination law in certain circumstances.

Psychological Injury and Stigma

Perhaps the preceding argument overlooks other features of discriminatory acts that always prove harmful, that adversely affect victims of discrimination regardless of their prior levels of well-being. More specifically, in maintaining Brest's earlier focus on the denial of opportunity discrimination inevitably entails, the preceding argument might seem to overlook discrimination's ability to harm its victims' psyches, injury that can afflict the rich as well as the poor, the powerful as well as the powerless, the socially integrated as well as the socially isolated. As Brest himself notes, "[d]ecisions based on assumptions of intrinsic worth and selective indifference inflict psychological injury by

stigmatizing their victims as inferior,”⁷⁴ and it seems we have little reason to think that such harm distinguishes among its victims according to their material circumstances. Unlike the denial of an opportunity, then, both psychological injury and stigma appear to cause harm in all contexts. Neither harm, moreover, presupposes perfectionism, since neither diverts our attention from the effects of discrimination to the moral wrongfulness of discriminatory acts or the moral characters of discriminators.

On closer inspection, however, it becomes clear that discrimination need not psychologically injure or stigmatize its victims nearly so frequently as we initially might suppose. Indeed, in many contexts, discriminators will inflict neither harm. For unlike the racists of the Jim Crow era, who often exploited instances of discrimination to insult and degrade their victims, most employers and purveyors of public accommodations who discriminate at all do so furtively, leaving their victims at best uncertain and at worst unaware of any malfeasance. A white racist employer, for example, will typically refrain from flaunting his bigotry in rejecting a black job applicant, but rather will contrive some ostensibly race-neutral rationale to dissemble his wrongdoing. Although the applicant might occasionally suspect the true, discriminatory nature of the employer’s decision, the pretext will often seem reasonable, obfuscating the discrimination for both the applicant and outside observers. Shrouded in such ambiguity, the discrimination is unlikely to cause harm through psychological injury or stigma, for the applicant will suffer no mental distress (beyond any disappointment stemming from the mere fact of the rejection itself), and his family, friends, neighbors, and fellow citizens will have no reason to view him as somehow tainted or inferior. An account that attempts to identify the harm of discrimination based on race or sex with psychological injury or stigma thus fails for the same reason Feinberg’s did: at least in certain circumstances, it cannot condemn a

⁷⁴ Brest, “Antidiscrimination Principle,” 8. For classic accounts of the harm caused by racial stigma, see Erving Goffman, *Stigma* (New York: Simon & Schuster, Inc., 1963); and Glenn C. Loury, *The Anatomy of Racial Inequality* (Cambridge, Mass.: Harvard University Press, 2003), ch. 3.

sufficiently large proportion of instances of discrimination to legitimate a continuous, categorical antidiscrimination regime.

Even if discrimination always psychologically injured and stigmatized its victims in all circumstances, the harm principle would not necessarily legitimize antidiscrimination law. For, as Gardner rightly notes, each harm is “fatally overinclusive.”⁷⁵ That is, we regularly suffer psychological injury—disappointment, distress, angst, and so on—in myriad contexts, yet liberals refuse to deem such injury harmful in the sense relevant for the harm principle. Similarly, while people often stigmatize others illegitimately, liberals do not permit the government to remedy all such wrongdoing. To pronounce the psychological injury and stigma wrought by discrimination harmful, then, would be to subject to governmental regulation other forms of objectionable behavior that liberals believe the government should tolerate.

We might think that the harm principle could avert this illiberal result, yet still legitimize antidiscrimination law, simply by distinguishing between legitimate and illegitimate forms of psychological injury and stigma, between those forms of psychological injury and stigma whose victims are *justified* in feeling disappointed or inferior and those forms whose victims should turn the other cheek—or, at least, not expect the government to intervene to protect them from unpleasant feelings. On this view, discrimination is one member of a small class of wrongdoing whose psychological effects are sufficiently objectionable to warrant governmental regulation.

But for the same reasons that Feinberg’s version of the harm principle could not cabin its legitimating force by distinguishing between legitimate and illegitimate setbacks to interests, the harm principle cannot do so by distinguishing between legitimate and illegitimate forms of psychological injury and stigma. Once again, in attempting to draw

⁷⁵ Gardner, “Liberals,” 6.

such a distinction, the harm principle faces a dilemma: either the basis for distinction is a perfectionist moral principle grounded in a conception of the good life, in which case the supplementary principle would compromise the harm principle's antiperfectionism, or the basis for distinction is an antiperfectionist moral principle, in which case the supplementary principle would efface the harm principle as a doctrine of political legitimacy altogether. If, for example, we supposed that people enjoy a Kantian right not to be insulted, humiliated, or stigmatized based on their race or sex, but that they enjoyed no similar right not to be insulted, humiliated, or stigmatized based on other factors, then the harm principle would no longer function as a doctrine of political legitimacy; the doctrine of Kantian rights, not the harm principle, would define the legitimate scope of governmental power, promulgating rights protection, not harm prevention, as the legitimate end of governmental regulation. Liberals cannot maintain the harm principle as a distinct antiperfectionist doctrine of political legitimacy and simultaneously appeal to extraneous moral principles to distinguish among different kinds of psychological and stigmatic effects.

For a similar reason, the harm principle cannot legitimize antidiscrimination law even if it turns out that discrimination need not psychologically injure or stigmatize its victims on every occasion in order to be harmful. Perhaps racial minorities and women can suffer psychological injury and stigma not only from discrete discriminatory acts, but also from the broader, systemic effects of such acts, taken in the aggregate. More specifically, were the government not to proscribe discrimination, racial minorities and women might be psychologically injured or stigmatized by the knowledge that other members of their racial group or sex were being discriminated against and that the government was tolerating the wrongdoing. They might, in short, come to view themselves as second-class citizens.

This account, of course, faces the difficult empirical question of how pervasive discrimination must be before racial minorities and women, as social groups, will indeed suffer psychological injury or stigma. But more fundamentally, the harm principle cannot recognize a harm that stems from people's psychological reactions to the government's failure to regulate a given form of conduct without compromising its antiperfectionism. For as we saw in Section II, many people have such reactions to the government's failure to regulate myriad kinds of conduct, including kinds of conduct that liberals believe the government should have to tolerate. For example, many people are psychologically injured or stigmatized by insults against members of their religion and the government's failure to proscribe such insults, perhaps even coming to view themselves as second-class citizens. And yet, most liberals understand the harm principle to require the government to tolerate such wrongdoing. To distinguish this tolerable wrongdoing from discrimination, the harm principle would have to distinguish between people's psychological reactions to the two kinds of conduct, a task it could perform only by advertent either to a conception of the good that would reduce the harm principle to a version of legal moralism or a doctrine of rights that would efface the harm principle as a doctrine of political legitimacy altogether. Neither option would preserve the harm principle as the antiperfectionist doctrine of political legitimacy that liberals have traditionally supposed it to be.

Harm to Dignity

Attempts to identify the harm of discrimination based on race or sex with psychological injury or stigma begin with an important insight: we condemn discrimination not simply for the adverse material consequences it has for its victims, but also for its other objectionable features. But such attempts nonetheless fail to legitimate antidiscrimination law according to the harm principle because they identify as those

other features contingent side-effects of discriminatory acts, subjective emotional reactions elicited by discrimination in some contexts but not in others. More specifically, while discrimination will sometimes psychologically injure and stigmatize its victims, it will frequently leave them unscathed—especially in contexts where discriminators conceal their wrongdoing. Therefore, if we are to pronounce discrimination harmful in all its instantiations, if we are to legitimize antidiscrimination law according to the harm principle as readily in a reasonably tolerant society as we could under apartheid, we must identify a feature of discriminatory acts that causes harm regardless of various contingencies, such as the material well-being of the victims or their respective vulnerabilities to psychological injury and stigma.

According to Denise Réaume, discrimination based on race or sex does indeed possess a feature that proves harmful in all contexts and thus always warrants governmental regulation according to the harm principle. In particular, Réaume argues that discrimination necessarily harms its victims’ “dignity,” impugning their very status as human beings whether they are affluent or indigent, stoic or flappable. All persons, Réaume contends, possess dignity in virtue of their “humanity,” their worthiness of “basic human respect.”⁷⁶ Irrespective of their personal qualities and worldly accomplishments, in other words, all persons deserve to be treated “as creatures of intrinsic, incomparable, and indelible worth, simply as human beings; no further qualifications necessary.”⁷⁷ Persons merit this treatment because they enjoy “self-determination or autonomy,” a capacity to exercise “agency” by making choices and charting their own course through life.⁷⁸ To respect other people’s dignity, we must recognize them as endowed with this capacity, treating them as sovereign beings rather than as mere automatons.

⁷⁶ Réaume, “Harm and Fault,” 361.

⁷⁷ Denise G. Réaume, “Discrimination and Dignity,” *Louisiana Law Review* 63 (2003): 675.

⁷⁸ Réaume, “Harm and Fault,” 370; Réaume, “Dignity,” 673.

Réaume, however, does not simply reduce autonomy to an objective, abstract capacity to choose. On the contrary, she believes that our autonomy comprises not only our ability to form a “conception of the good,” but also our ability to form a “conception of the self.”⁷⁹ To enjoy dignity, then, a person must be able to embrace his or her “identity” or “sense of self,” conceiving of himself or herself as “a self-conscious being with a secure sense of his or her worth and place in the world, in command of his or her life and under no one’s thumb.”⁸⁰ One will not develop this self-conception, in turn, unless others respect “those features of oneself that are important elements of identity” and in which one “is entitled to take pride.”⁸¹ To be sure, different personal “features” will constitute “important elements of identity” in different social contexts, for we construct our identities partially by reference to social norms and practices.⁸² But because we will always attach special importance to some of our personal traits, because we will always regard our agency as inextricably intertwined with our “sense of self,” we will enjoy dignity and claim our equal standing among our fellow human beings only if we can view our identities with equanimity.

One violates other people’s dignity, Réaume argues, if one treats others in ways that are inconsistent with their humanity. More specifically, in violating others’ dignity, one “convey[s] the message that some are of lesser worth than others,” that they lack intrinsic value and thus do not merit respect.⁸³ According to Réaume, this kind of treatment not only wrongs its victims, but also harms them, setting back their interests as readily as would a denial of opportunity or psychological injury or stigma. To a much greater extent than these other harms, however, “harm to dignity” derives its force from context rather than consequences, from the “social meaning of one’s treatment by others”

⁷⁹ Réaume, “Dignity,” 676-77.

⁸⁰ *Ibid.*, 674.

⁸¹ Réaume, “Harm and Fault,” 375, 361.

⁸² *Ibid.*, 361-62.

⁸³ Réaume, “Dignity,” 672, 675-76.

rather than the treatment's "tangible effects."⁸⁴ We thus cannot determine precisely what kind of treatment will harm persons' dignity without engaging in an "interpretive exercise" that seeks to identify the "objective or social meaning of certain forms of behavior."⁸⁵ Once we have grasped the "objective or social meaning" of our conduct, though, we can comport ourselves so as to avoid harming others by implying their inferior moral worth.

Although Réaume seems at times to flirt with a perfectionist understanding of "harm to dignity," she ultimately disclaims any concern with moral character, insisting that the harm inheres not in the motivations of the perpetrator, but solely in the effects on the victim. Neither material nor psychological, these effects amount to the denial of the victim's humanity, a denial whose meaning is determined by social context rather than individual motives. Consequently, even if one in no way intends such a result, one can harm another person's dignity so long as social convention imbues one's conduct with a disrespectful connotation.⁸⁶ Because even the virtuous can thus disrespect their fellow human beings, we need not appraise persons' moral characters to identify instances of "harm to dignity."

To be sure, "harm to dignity" will be most flagrant in cases where the person causing the harm does indeed act from a vicious character, consciously intending to impugn his victim's moral worth. It is precisely such a desire that seems to motivate genuine racists and sexists who discriminate directly against racial minorities and women. When, for example, an employer refuses to hire an applicant "out of racial or other *animus* or out of a belief in her inferiority," he harms her dignity by "treat[ing] her as unworthy of basic human respect," as less than a full human being.⁸⁷ More specifically, the employer deliberately denigrates an aspect of the applicant's identity, perpetrating "an

⁸⁴ Réaume, "Harm and Fault," 368.

⁸⁵ Réaume, "Dignity," 684, 680.

⁸⁶ Réaume, "Harm and Fault," 365-66.

⁸⁷ *Ibid.*, 361.

assault on [her] sense of self” and hence her status as an autonomous agent.⁸⁸ As we have seen, this harm reduces to neither the material nor the psychological effects of the discrimination for the victim, but is instead distinct from both the denial of opportunity and the “hurt feelings” she suffers.⁸⁹ The most paradigmatic applications of antidiscrimination law, Réaume thinks, attempt to remedy the independent “harm to dignity” wrought by particularly egregious instances of direct discrimination.

But because social context largely determines the meaning our conduct has for others, a discriminator need not espouse racist or sexist beliefs for his direct discrimination to cause “harm to dignity.” Consider, for example, the white restaurateur who refuses to serve black people not because he himself believes they are morally inferior, but because his predominantly white, racist clientele would refuse to patronize a racially integrated establishment. Although the restaurateur discriminates “as a means to another end” rather than “as an end in itself,” he harms his victims’ dignity nonetheless.⁹⁰ For in accommodating his white customers’ racist preferences, the restaurateur engages in conduct whose social meaning amounts to a tacit denial of black people’s intrinsic moral worth, a kind of “second-hand insult” that affronts dignity as piercingly as does blatant racism.⁹¹ According to Réaume, antidiscrimination law comprehends this more oblique harm by proscribing direct discrimination irrespective of its perpetrators’ motives.

Indeed, Réaume suggests that even antidiscrimination law’s prohibitions against indirect discrimination seek to remedy “harm to dignity.” When, for instance, an employer refuses to hire anyone under a particular height (say, five feet, ten inches) for a job, he inevitably curtails opportunities disproportionately for women, who tend, on average, to be significantly shorter than men. To be sure, the height requirement will be

⁸⁸ Réaume, “Dignity,” 679.

⁸⁹ Réaume, “Harm and Fault,” 361, 362.

⁹⁰ *Ibid.*, 366.

⁹¹ *Ibid.*, 367.

perfectly legitimate if one truly must be tall to perform the job successfully, a possibility the antidiscrimination regimes of most jurisdictions recognize.⁹² But if a short person can perform the job as well as a tall person, then the height requirement serves as a mere proxy for sex, arbitrarily excluding perfectly qualified applicants from consideration. According to Réaume, such groundless requirements “reify the attributes of the typical incumbent employee, thereby creating a standard image of the job’s model occupant and building it into the job description.”⁹³ Those who do not conform to this “standard image,” in turn, suffer “harm to dignity,” for the exclusionary job description tacitly impugns an important aspect of their identity and thus devalues their agency.⁹⁴ Antidiscrimination law attempts to forestall this harm by prohibiting facially neutral but practically discriminatory job requirements.

Direct or indirect, overt or subtle—discrimination based on race or sex seems to violate its victims’ dignity in all circumstances. But is Réaume correct to insist that these violations of dignity are *harmful* in the sense relevant for the harm principle? Even if such violations are profoundly immoral, as they undoubtedly are, the harm principle will license the government to regulate them only if they also cause harm, only if they diminish their victims’ life prospects. Réaume’s use of the term “harm to dignity” notwithstanding, violations of dignity need not—and, indeed, frequently do not—harm their victims. Réaume renders the opposite conclusion plausible only by adulterating her otherwise objective account of violations of dignity with subjective elements, elements that evoke the putative harms of psychological injury and stigma considered above. As we discovered, however, neither harm can distinguish discrimination from myriad other wrongs the harm principle forbids the government to regulate. If, by contrast, we construe “harm to dignity” as a purely objective wrong, one completely independent of the

⁹² See, e.g., Equality Act 2010, § 19.

⁹³ Réaume, “Harm and Fault,” 376.

⁹⁴ *Ibid.*, 377.

subjective mental states of its victims, we will still fail to satisfy the harm principle, since, on such an interpretation, it matters not what effects violations of dignity might have on the quality of their victims' lives, effects that are the sine qua non of any liberal understanding of the concept of harm. Réaume's account of the moral wrongfulness of discrimination thus faces a dilemma: either the wrong is a subjective one consisting in psychological injury and stigma, in which case discrimination does not differ significantly from other wrongs a liberal, antiperfectionist government may not regulate, or the wrong is a purely objective one that is in no way harmful. Whichever horn of the dilemma Réaume would embrace, her notion of "harm to dignity" cannot legitimate antidiscrimination law according to the harm principle.

Discerning Réaume's true understanding of that notion is rather difficult, for she continually equivocates between subjective and objective conceptions of violations of dignity. To be sure, Réaume repeatedly insists that such violations are purely objective harms rather than subjective, psychological ones. "Harm to dignity," Réaume avers in a typical passage, "must be understood to inhere in the denial of respect *per se*. In other words, harm to dignity is better understood as an independent, objective harm, not a matter of hurt feelings."⁹⁵ On this understanding, we must identify "harm to dignity" not through an "empirical" examination of the "psychological effects" violations of dignity might have for their victims, but through a "normative" analysis of the objective social meaning of different forms of conduct.⁹⁶ But Réaume herself frequently flouts this injunction, explicating the alleged harm of discrimination in terms of its subjective impact on its victims' psyches. By exposing its victims to "ridicule and abuse," Réaume explains, discrimination "stigmatizes and often humiliates"⁹⁷ and thus "can scarcely fail to be

⁹⁵ Ibid., 362. See also Réaume, "Dignity," 676, 682-83.

⁹⁶ Réaume, "Dignity," 684.

⁹⁷ Réaume, "Harm and Fault," 361, 361-62.

experienced as demeaning because it *is* demeaning.”⁹⁸ Réaume can validly conclude that such stigma and humiliation violate dignity only because she defines that value in terms of the subjective notions of “identity” and “self-esteem,” which, in turn, involve “being able to feel at home with those features of oneself that are important elements of identity.”⁹⁹ One’s dignity, it thus turns out, consists not simply in one’s intrinsic moral value, one’s worthiness of respect, but also in one’s beliefs about and attitudes toward one’s self and identity. When juxtaposed with this subjective conception of dignity, Réaume’s repeated assertions that violations of dignity are purely objective harms are simply incoherent.

Whether we decide to resolve this incoherence in favor of a subjective understanding of “harm to dignity” or an objective one, discrimination will not cause harm simply in virtue of its violating dignity. First, let us construe “harm to dignity” as a subjective wrong, as a form of “ridicule and abuse” that “stigmatizes and often humiliates” its victims, preventing them from “taking pride” in important aspects of their “identity” and hence impairing their “self-esteem.” Like the psychological and stigmatic accounts considered above, this subjective understanding of “harm to dignity” cannot distinguish discrimination from other types of objectionable behavior that the harm principle forbids the government to remedy and thus cannot legitimize antidiscrimination law.

Nor will espousing a purely objective understanding of “harm to dignity” rescue Réaume’s account of the harm of discrimination. For by deeming violations of dignity to be harmful *per se*, regardless of their material and psychological consequences for their victims, such an understanding distorts the concept of harm beyond all recognition. As Gardner and Shute rightly note, we ordinarily conceive of harms as “instrumental wrongs” that entail “the diminution of people’s prospects, the change for the worse in their

⁹⁸ Réaume, “Dignity,” 681.

⁹⁹ Réaume, “Harm and Fault,” 375.

lives.”¹⁰⁰ Réaume, by contrast, considers even “‘Kantian’, non-instrumental wrongs”¹⁰¹—like violations of dignity, understood objectively—to be harmful. Like Antony Duff, then, she recognizes a category of harmful behavior that engenders no corresponding “harmed condition” in its victims, causing harm not through its tangible effects, but through its social meaning.¹⁰² This position supposes that no wrongdoing is truly harmless, that a wrong always harms its victims “just by virtue of being a wrong.”¹⁰³ In so conflating harm with wrongdoing more generally, however, Réaume effectively annuls the harm principle’s protections of individual liberty. For if all wrongs are harmful by definition, if a wrongdoer cannot but cause harm through his wrongdoing, then the harm principle will license the government to regulate all wrongful behavior, placing no categorical limits on the exercise of political power. To pronounce all wrongs per se harmful thus is effectively to disavow the harm principle as a liberal doctrine of political legitimacy and to espouse legal moralism. In insisting that violations of dignity, construed as purely objective wrongs, do indeed cause harm, Réaume forsakes liberalism in precisely this way, but it seems few liberals are likely to follow suit.

And yet, we should not infer from this critique that Réaume’s theory of dignity fails as an account of the *moral wrongfulness* of discrimination. On the contrary, we seem to presuppose such a theory when we condemn discrimination as morally wrongful regardless of the consequences it has for its victims and advocate antidiscrimination law as fervently in tolerant societies as we would under Jim Crow. But moral wrongfulness is one thing, and harmfulness is quite another. The harm principle owes its *raison d’être* to the fact that it subjects to governmental regulation not all moral wrongs, but only those that also cause harm. Like many of the doctrine’s adherents, Réaume succumbs to the

¹⁰⁰ Gardner and Shute, “Rape,” 216.

¹⁰¹ Ibid.

¹⁰² R.A. Duff, “Harms and Wrongs,” *Buffalo Criminal Law Review* 5 (2001): 23-25.

¹⁰³ Ibid., 25-26.

temptation to disregard this constraint when confronted with a particularly flagrant, but nonetheless harmless, moral wrong, privileging her moral indignation over her commitment to liberalism.

V. CONCLUSION

Many liberals tend to assume that discrimination based on race or sex harms its victims and that the harm principle consequently legitimates antidiscrimination law. Considered against the backdrop of Jim Crow and apartheid, this assumption is indeed warranted, for such regimes of oppression denigrate their targets and deny them even the rudiments of a minimally decent life, fostering a climate of humiliation and deprivation in which nearly every instance of discrimination proves harmful. But as this chapter has argued, we can imagine a society that has ameliorated its moral culture to such an extent that the practice of discrimination often has no such deleterious effects. To be sure, even in such a hypothetical society, some women and racial minorities—particularly those consigned to the lower socioeconomic classes—will remain acutely vulnerable to the denials of opportunity entailed by discrimination, vulnerability that will support a limited harm-based rationale for antidiscrimination law. As applied to the more fortunate members of traditionally disadvantaged groups, however, antidiscrimination law will lose this rationale, serving merely to punish especially deplorable, but nonetheless harmless, wrongdoing. Only by expanding the concept of harm to encompass ever broader swaths of the moral terrain—or even, as in the case of “harm to dignity,” all of morality—can liberal proponents of the harm principle subject to governmental regulation even the less injurious instances of discrimination and thereby legitimize the continuous, categorical antidiscrimination regimes adopted by most contemporary western liberal democracies. To justify antidiscrimination law as we know it, liberal proponents of the harm principle must abjure their liberalism.

Some might object that this conclusion mischaracterizes liberalism, offering a tendentious caricature that exaggerates liberalism's reluctance to restrict individual liberty for the sake of other values and downplays its sensitivity to moral wrongs other than harm. As we saw in the Introduction, however, it is a version of liberalism to which many self-avowed liberals subscribe. For most contemporary liberals do indeed attach a nearly absolute priority to individual liberty over other values and discount moral wrongs that, unlike harm, cannot be identified without adverting to a conception of the good life. These features of contemporary liberalism owe largely to its antiperfectionism, which permits the government to proscribe only that wrongdoing we all can reasonably recognize, whatever moral doctrines we happen to espouse, and demands that such proscriptions be justified as measures not for improving persons' moral characters, but for forestalling and remedying the adverse consequences of individual behavior. Unless we are willing to relax these restrictions, we will have to abide many instances of discrimination based on race or sex.

Chapter 4

Liberal Perfectionism

I. INDIVIDUALISTIC WRONGS AND LIBERAL PERFECTIONISM

Antiperfectionist liberal doctrines of political legitimacy fail to justify many core applications of antidiscrimination law. Attempting to eschew evaluations of individual moral character and thus to heed the strictures of public reason, many such doctrines adopt a systemic perspective, construing discrimination based on race or sex not as a unique injustice consisting in discrete discriminatory acts, but as merely one among many manifestations of a corrupt institutional structure that systematically disadvantages particular social groups. As Chapter 2 demonstrated, however, such an account of discrimination renders much antidiscrimination law illegitimate. For in considering discrimination only for its broader, systemic effects, an antiperfectionist liberal theory must condone those instances of discrimination whose regulation would infringe the individual rights and liberties with which antiperfectionist liberalism limits the pursuit of systemic goals. To condemn even these instances of discrimination and thereby legitimize all applications of antidiscrimination law, we would have to adopt an individualistic perspective and develop a fuller account of the objectionable nature of discrete discriminatory acts.

Other liberal doctrines of political legitimacy promise precisely such an account of discrimination while nevertheless disclaiming any allegiance to a particular conception of the good life and thus remaining thoroughly antiperfectionist. Most notably, by focusing exclusively on the deleterious effects discrimination has for its individual victims, an antiperfectionist version of the harm principle purports to legitimate all applications of antidiscrimination law without having to adjudge the morality of any

action or the quality of anyone's character. But this claim turns out to be spurious, since, as we discovered in Chapter 3, not all instances of discrimination harm their victims.

Only by compromising either its liberalism or its antiperfectionism could an individualistic account of discrimination like that entailed by the harm principle condemn all discriminatory acts and legitimize all applications of antidiscrimination law.

Indeed, to legitimize the full extent of modern antidiscrimination regimes in all circumstances, it seems a doctrine of political legitimacy must comprehend nearly all of morality, adverting to a conception of the good life, flouting public reason, and forsaking antiperfectionism for perfectionism. For unless the doctrine is willing to recognize at least some moral wrongs that cannot be explicated in purely antiperfectionist terms, it will so circumscribe political power as to place many instances of discrimination beyond the legitimate scope of governmental regulation. Liberals must choose between their commitment to antidiscrimination law and their aversion to legal moralism.

Because they assume that perfectionism and liberalism are radically incompatible, many liberals will balk at embracing legal moralism to vindicate antidiscrimination law. And yet, perhaps a doctrine of political legitimacy can license perfectionist regulation, regulation that attends to the morality of individual behavior and the characters of the persons who engage in it, while nonetheless honoring the fundamental tenets of liberalism identified in the Introduction. Recall that underlying liberalism is a conception of the person as an autonomous being, a sovereign agent who must decide for himself what conception of the good life he will espouse and under what social institutions and rules he will live. Liberals derive from this conception of the person an ideal of public justification according to which we must defend our common social institutions and practices to one another in terms we all can reasonably accept, whatever our personal moral views.

Do perfectionist doctrines of political legitimacy necessarily violate this condition and thereby render themselves illiberal? Or can perfectionist doctrines endorse particular conceptions of the good life while espousing liberalism's conception of the person and its concomitant ideal of public justification? Contemporary liberals' misgivings notwithstanding, no inherent feature of such doctrines precludes their conceiving of persons as autonomous beings. For many conceptions of the good life prize personal autonomy, deeming the value an essential prerequisite for or ingredient of complete human flourishing.

Liberalism's ideal of public justification, by contrast, precludes perfectionism. At least when construed strictly to purge the public arena of all consideration of conceptions of the good life, that ideal contravenes the animating rationale of perfectionist doctrines of political legitimacy, which license the government to promote morality and cultivate persons' moral characters. In particular, because we fundamentally disagree about precisely which moral principles the government should promote and precisely which character traits it should cultivate, many liberals suppose that the government must remain neutral among competing conceptions of the good and ground its policies only in more ecumenical principles, a requirement that is anathema to the perfectionist project.

But what if the government espouses a distinctively liberal version of perfectionism, one that seeks not to impose on us any single, monolithic moral vision, but rather to help us to navigate a pluralistic moral reality according to our own, autonomous choices? While a government that pursues the perfectionist project, so defined, will violate the strict interpretation of the liberal ideal of public justification, extolling certain ways of life as valuable and deprecating others as worthless, it will nevertheless countenance enough activities and relationships that it can avoid systematically favoring the beliefs of some of its citizens over those of others. Moreover, because such a

government respects personal autonomy, it will promote morality and inculcate virtue only through means consistent with that value, thus forgoing those means, like coercion, that liberals find most repugnant. It thus seems that a perfectionist doctrine of political legitimacy that esteems autonomy and acknowledges value pluralism can accord at least with the spirit of liberalism's foundational tenets.¹

Joseph Raz attempts to reconcile perfectionism and liberalism in precisely this way, charging the government with the perfectionist mission of fostering certain conceptions of the good life, yet maintaining a due regard for value pluralism and personal autonomy.² In so alloying perfectionism with fundamental liberal principles, Raz seems to have forged just the kind of liberal doctrine of political legitimacy needed to legitimate all applications of antidiscrimination law in contemporary circumstances. For the antiperfectionist doctrines considered in the preceding two chapters foundered on their compunction against acknowledging certain classes of moral wrongs and evaluating moral character, a compunction that liberal-perfectionist doctrines like Raz's do not share. Because it sheds the fetters of antiperfectionism, because it permits the government—at least in principle—to remedy all moral wrongs, Raz's doctrine of political legitimacy can apparently condemn all instances of discrimination.

Liberalism has a vital stake in the success of Raz's project. Having failed to justify antidiscrimination law in antiperfectionist terms, liberalism must now turn to perfectionism for a more robust account of the wrongfulness of discrimination—without compromising its own fundamental commitments in the process. Only a theory like Raz's seems to permit liberalism to strike this delicate balance between a morally partisan political agenda, on the one hand, and a categorical commitment to personal autonomy,

¹ For recent works that deny the compatibility of liberalism and perfectionism, see Martha C. Nussbaum, "Perfectionist Liberalism and Political Liberalism," *Philosophy & Public Affairs* 39 (2011): 3-45; Jonathan Quong, *Liberalism Without Perfection* (Oxford: Oxford University Press, 2011), esp. ch. 2.

² See Joseph Raz, *The Morality of Freedom* (Oxford: Oxford University Press, 1986), chs. 14 and 15.

on the other. But if such a liberal-perfectionist union proves untenable, if, in particular, the government can rectify all instances of discrimination only by occasionally subordinating personal autonomy to legal moralism, then it seems we will have to look beyond liberalism and toward unadulterated perfectionism for a comprehensive justification of antidiscrimination law.

This chapter contends that Raz's liberal perfectionism does indeed fail to legitimate many applications of antidiscrimination law. In fact, his favored doctrine of political legitimacy, a perfectionist reconstruction of the liberal harm principle, is hamstrung by limitations similar to the ones that bedeviled its antiperfectionist cousin in the previous chapter. To be sure, unlike an antiperfectionist version of the harm principle, Raz's may invoke supplementary moral principles to distinguish between legitimate and illegitimate harms and thereby cabin its own legitimating force. But like the antiperfectionist accounts of harm considered in the previous chapter, Raz's autonomy-based account fails to pronounce all instances of discrimination harmful, for not every discriminatory act violates its victim's autonomy, as that concept is understood by Raz. Moreover, by exalting personal autonomy as the paramount value for a perfectionist government to promote, Raz actually compounds the antiperfectionist harm principle's shortcomings, affording discriminatory conduct even more protection from governmental interference. Insofar as it incorporates liberalism's fundamental tenets, Raz's perfectionism also assumes some of the ideology's liabilities—in particular, an incapacity to legitimize certain core applications of antidiscrimination law.

II. RAZ'S PERFECTIONIST HARM PRINCIPLE

At least on initial inspection, Raz's perfectionist doctrine of political legitimacy bears little resemblance to the antiperfectionist ones considered in the previous two chapters. Indeed, Raz explicitly repudiates the two possible theoretical bases for those

other doctrines, unequivocally denying that the government must remain neutral among competing conceptions of the good life³ or abstain from acting on reasons derived from such conceptions.⁴ Were the government to heed these calls for neutrality or abstentionism, Raz argues, it would thwart the very purposes it is supposed to serve. Whereas antiperfectionists would have the government perform only some relatively limited function like securing distributive justice or guaranteeing individual rights, Raz endows it with an expansive writ, insisting that “[t]he main purpose of government . . . is to assist people . . . to lead successful and fulfilling lives, or, to put the same point in other terms, to protect and promote the well-being of people.”⁵ And because Raz believes that people’s well-being consists in the successful pursuit only of *valuable* goals, the government’s duty to promote people’s well-being entails a subsidiary duty “to promote morality,” to “promote the moral quality of the life of those whose lives and actions [it] can affect.”⁶ The government can discharge these duties only if it may consider the full range of reasons relevant to its various policy decisions—including reasons derived from a conception of the good life. Against antiperfectionists, then, Raz deems all moral reasons “fair game for governmental action,” imposing “no fundamental principled inhibition on governments acting for any valid moral reason”⁷ and thus spurning the liberal commitment to public reason.

Once we consider the content of the conception of the good Raz charges the government with promoting, however, we discover that his perfectionism is much more congenial to liberalism’s foundational tenets than it initially appears to be. For Raz founds that conception on the value of personal autonomy, a value that, as we have seen, even

³ See *ibid.*, ch. 5.

⁴ See *ibid.*, ch. 6.

⁵ Joseph Raz, “Liberty and Trust,” in *Natural Law, Liberalism, and Morality*, ed. Robert P. George (Oxford: Oxford University Press, 1996), 113.

⁶ Raz, *Morality of Freedom*, 415.

⁷ Joseph Raz, “Facing up: A Reply,” *Southern California Law Review* 62 (1989): 1230.

antiperfectionist liberals champion. At least amid the post-industrial conditions prevailing in contemporary western liberal democracies,⁸ Raz contends, a person cannot flourish, a person cannot lead a successful life, unless he enjoys and exercises autonomy, freely choosing for himself what goals, projects, and relationships to pursue; hence our tendency to evaluate the success of a person's life "not only by the success of his projects but also by how he came to have them."⁹ We suppose, in other words, that, in realizing a valuable goal, one enhances one's own well-being only if one has freely chosen to pursue the goal, only if one realizes the goal autonomously. By thus pronouncing autonomy "an essential ingredient of individual well-being" and "a constituent element of the good life,"¹⁰ Raz injects into his conception of the good a heavy dose of liberal dogma.

Raz only confirms these affinities between his conception of the good and fundamental liberal tenets when he expounds his ideal of personal autonomy, an ideal that seems to correspond rather closely to the liberal conception of the person. According to that ideal, "people should make their own lives."¹¹ Raz thus conceives of autonomous persons as "those who can shape their life and determine its course," "part creators of their own moral world" who "control[], to some degree, their own destiny, fashioning it through successive decisions throughout their lives."¹² Unlike "one who drifts through life unawares," the autonomous person recognizes that he faces myriad options and that he partially structures his existence by choosing among them.¹³ Not a life of "coerced choices," of "no choices," or of simply "drifting through life without ever exercising

⁸ Raz, *Morality of Freedom*, 394. Raz's argument for the value of autonomy is thus "contextualist" in that its validity presupposes the existence of certain social conditions. For a critique of this argument, see David McCabe, "Joseph Raz and the Contextual Argument for Liberal Perfectionism," *Ethics* 111 (2001): 493-522.

⁹ Raz, *Morality of Freedom*, 377.

¹⁰ *Ibid.*, 369, 408.

¹¹ *Ibid.*, 369.

¹² *Ibid.*, 154, 369.

¹³ *Ibid.*, 381-82.

one's capacity to choose,"¹⁴ but a life of freely chosen pursuits—that is the life Raz's ideal of autonomy, like the liberal conception of the person, extols and aims to promote.

Raz's ideal of autonomy, however, diverges from antiperfectionist ones in refusing to pronounce personal autonomy categorically valuable, valuable irrespective of the moral quality of the goals its possessors pursue. Instead, Raz insists that "[a]utonomy has no value *qua* autonomy when it is abused";¹⁵ a person promotes his well-being through autonomous choices only when he exercises his autonomy "in pursuit of the good," only, that is, when he chooses to realize morally valuable options.¹⁶ In so conditioning the value of autonomy on the value of the goals pursued through autonomous choices, Raz unequivocally proclaims his perfectionism.

Were Raz also to espouse a particularly narrow account of the values to be instantiated in autonomous choices, an account that deemed only a few activities and relationships valuable and thus worthy of pursuing, he would alienate his theory not only from antiperfectionism, but also from liberalism more generally. For whether perfectionist or antiperfectionist, most liberals consider the diversity of moral beliefs and ways of life that characterizes contemporary western liberal democracies to be not just some accident of our social development, but rather a reflection of a deeper moral reality.¹⁷ But Raz himself recognizes and embraces this reality, contending that one cannot value autonomy without also endorsing the thesis of "value pluralism."¹⁸ According to this thesis, "there are various forms and styles of life which exemplify different virtues and which are incompatible," which "cannot normally be exemplified in the same life."¹⁹ To be sure, we must not conflate value pluralism, so defined, with the

¹⁴ Ibid., 371.

¹⁵ Ibid., 380.

¹⁶ Ibid., 381.

¹⁷ See, e.g., Isaiah Berlin, *Four Essays on Liberty* (Oxford: Oxford University Press, 1990).

¹⁸ Raz, *Morality of Freedom*, 398.

¹⁹ Ibid., 395.

antiperfectionist's neutrality among competing conceptions of the good life, for unlike the latter, the former "allows that certain conceptions of the good are worthless and demeaning, and that political action may and should be taken to eradicate or at least curtail them."²⁰ Despite this discrimination among conceptions of the good, however, Raz's value pluralism encompasses a sufficiently broad range of activities and relationships to share liberalism's respect for lifestyle diversity and its concomitant wariness of governmental moralizing.

From this attempted reconciliation of Raz's perfectionism with liberalism, we should not conclude that Raz initially empowers the government to promote morality only to annul that authority by espousing a vacuous conception of the good, a conception so catholic in the activities and relationships it deems valuable that it obviates the need for the government to guide people toward the good life. On the contrary, Raz insists that, even while respecting individual choice and acknowledging value pluralism, the government must assume an active role in securing the prerequisites of autonomy, a role that reflects the duty each of us has to promote autonomy. For if autonomy is in fact valuable, if one cannot flourish without living autonomously, then we each have reason to make not only ourselves, but also all other human beings, autonomous.²¹ To be sure, neither the government nor any individual person can literally *make* someone else autonomous, since the decision to live autonomously must itself be made freely.²² But although we cannot compel our fellow human beings actually to live autonomous lives, we nevertheless can help them to develop the "capacity" for autonomy, the combination of personal faculties and life circumstances one must enjoy if any of one's choices are to be autonomous.²³ More specifically, one cannot be truly autonomous unless one has

²⁰ Ibid., 133.

²¹ Ibid., 407.

²² Ibid.

²³ Ibid., 372. See also *ibid.*, 204.

“appropriate mental abilities, an adequate range of options, and independence [i.e., freedom from coercion and manipulation].”²⁴ Because only the government can ensure that these conditions obtain broadly across society, it must attempt to foster them through its policies—not only by refraining from unduly coercing and manipulating its citizens, but also by furnishing them with various resources and opportunities.²⁵ These autonomy-based duties thus vest the government with the perfectionist mission of facilitating its citizens’ pursuit of the good life, a mission rendered no less moralistic by the fact of value pluralism.

Nor does Raz restrict the government’s moralizing merely to collective measures aimed at securing the preconditions of autonomy. An autonomy-promoting government like the one Raz envisages also must ensure that its citizens discharge their own, individual autonomy-based duties in their regular interactions with one another, for autonomy presupposes not only a hospitable general milieu, but also particular interpersonal relationships of a certain character. If people generally shirked their autonomy-based duties, either impairing others’ capacities for autonomy or failing affirmatively to help others to cultivate those capacities, few people would lead autonomous lives—regardless of any other autonomy-promoting policies the government might enact. Raz’s perfectionism thus authorizes the government “to use coercion both in order to stop people from actions which would diminish people’s autonomy and in order to force them to take actions which are required to improve peoples’ [*sic*] options and opportunities.”²⁶ Raz’s perfectionism, in other words, assigns the government precisely the role we would expect a perfectionist government to perform: helping its citizens to live moral lives by, *inter alia*, compelling them to honor the duties whose fulfillment is

²⁴ *Ibid.*, 372-73.

²⁵ Raz, “Liberty and Trust,” 113.

²⁶ Raz, *Morality of Freedom*, 416.

necessary to support a social milieu that is conducive to a particular conception of the good life.

But in what sense does Raz remain a liberal? Were Raz to authorize the government to enforce *all* its citizens' moral obligations—and by any means at its disposal, subject only to considerations of efficacy—the two visions would indeed be hopelessly irreconcilable. But in fact, Raz assigns the government a much narrower brief, forbidding it to employ the power of the law (though not necessarily other means) to compel its citizens to fulfill their non-autonomy-based duties. More specifically, like many contemporary liberals, Raz endorses the harm principle as the appropriate doctrine of political legitimacy and tethers the concept of harm to that of autonomy so that only violations of autonomy-based duties qualify as harmful and thus amenable to legal regulation. By discounting violations of other, non-autonomy-based moral duties as harmless, Raz significantly curtails the government's power, tempering his perfectionism and burnishing his liberal bona fides.

Perhaps somewhat paradoxically, Raz is led to this quintessentially liberal position by his perfectionist commitment to personal autonomy. For that commitment spawns a sweeping suspicion of coercion in general, which, in turn, entails a more targeted skepticism of governmental coercion in particular. As noted above, Raz believes that one must enjoy “independence,” or freedom from coercion (and manipulation²⁷), to be truly autonomous.²⁸ Roughly put, one person, *A*, coerces another, *B*, not to perform a given action, *Z*, when *A* deliberately and wrongfully induces *B* to forgo *Z* by rendering the consequences for *B* of performing *Z* so undesirable that *B* comes to regard *Z* as

²⁷ The discussion below ignores the issue of manipulation both because Raz's own treatment of it is so cursory and because coercion, not manipulation, is traditionally taken by liberals to be the harm principle's primary concern.

²⁸ Raz, *Morality of Freedom*, 373.

unchoiceworthy, as no longer eligible as an option for him to pursue.²⁹ Such treatment compromises its victims' autonomy not only by effectively depriving them of options, but also by subjecting them to other people's wills and thus reducing them to mere instruments to be employed for others' purposes.³⁰ Therefore, to respect properly the value of autonomy, to discharge fully our autonomy-based duties, we must refrain from illegitimately coercing others.³¹

Itself a moral agent that can affect people's autonomy, the government, too, must observe the autonomy-based duty not to coerce people illegitimately. Indeed, the government has a particular penchant for impairing people's autonomy through coercion, since it can render law, the distinctive means of governmental action, effective only by credibly threatening to coerce those recalcitrant people who flout its commands. But in acting on this threat and enforcing the law against its citizens, the government necessarily curtails autonomy. For given the "symbolic meaning" coercion has garnered in our society, the coercive enforcement of the law "expresses a relation of domination and an attitude of disrespect for the coerced individual," depreciating his social standing below the level at which he can command from his fellow citizens the regard befitting an autonomous agent. In addition, legal regulation is inevitably "indiscriminate": even when it aims to eliminate only immoral options, it will inadvertently extinguish valuable ones as well, a constriction of choice that diminishes autonomy.³² The government can avert these harms to autonomy and honor its autonomy-based duties only if it eschews coercion except under very limited conditions—namely, when it can demonstrate that coercion will promote more autonomy than it curbs.

²⁹ For Raz's more precise account of coercion, see *ibid.*, 148-49.

³⁰ *Ibid.*, 155-57, 410.

³¹ *Ibid.*, 407.

³² *Ibid.*, 418-19.

The government will meet this burden, Raz contends, only when it can satisfy the criterion of political legitimacy embodied in the liberal harm principle. Raz believes, in other words, that “autonomy-based duties never justify coercion where there is no harm,” where no one’s autonomy has been compromised, and thus “regards the prevention of harm to anyone (himself included) as the only justifiable ground for coercive interference with a person.”³³ When the government employs coercion not to forestall harm to autonomy, but to achieve other ends, it restricts some people’s autonomy without realizing any compensating benefits for the autonomy of others (or even for the long-term autonomy of the coerced people themselves).³⁴ A version of the harm principle rooted in a commitment to autonomy condemns such instances of governmental coercion. By endorsing the harm principle as the proper doctrine of political legitimacy, then, Raz significantly tempers his perfectionism, denying officials who seek to promote non-autonomy-based moral principles and to cultivate non-autonomy-based moral virtues recourse to one of the government’s most potent implements: the coercive power of the law.

Lest we fear that he has now forsaken perfectionism altogether, espousing the very neutrality and abstentionism he earlier repudiated, Raz emphasizes that his version of the harm principle still affords the government much latitude to promote morality and cultivate character. For unlike the antiperfectionist versions of the harm principle encountered in the previous chapter, which exalted the doctrine as a “fundamental principle of morality,” Raz’s perfectionist version attempts “to downgrade it to a sound practical principle to guide governmental action.”³⁵ Raz conceives of the harm principle, in other words, “not as a restraint on the pursuit of moral goals by the state, but as

³³ Ibid., 415, 412-13.

³⁴ Ibid., 419.

³⁵ Raz, “Facing up,” 1229 n.162.

indicating the right way in which the state could promote the well-being of people.”³⁶ His version of the doctrine thus permits the government to induce its citizens to live moral lives by, inter alia, “subsidiz[ing] certain activities, reward[ing] their pursuit, and adverti[sing] their availability.”³⁷ So long as the government abjures coercion—particularly legal coercion—as a means of achieving its most moralistic ends, it complies with the harm principle and fulfills one of its most important autonomy-based duties. Not an ends-limiting doctrine of political legitimacy, but one that “sets a limit on the means allowed in pursuit of moral ideals”³⁸—Raz’s harm principle fully accords with a broader perfectionist agenda.

At the same time, however, Raz’s embrace of the harm principle bespeaks the fundamentally liberal nature of his perfectionism. For unlike non-liberal perfectionist theories, Raz’s perfectionist one considers personal autonomy to be not merely one important value among others, but the paramount value for political morality, the value that largely determines the government’s policy agenda and demarcates its limits. This uncompromising commitment to autonomy distinguishes liberalism from its ideological competitors and yields principles of political morality—most notably, the harm principle itself—that are inconsistent with those of non-liberal versions of perfectionism. In particular, “non-liberal theories that nevertheless value individual liberty and embrace pluralism” reject the “liberal idea that there are strict moral norms (and not merely prudential limits) that exclude in principle moral paternalism and the use of coercion to prevent moral harm” and countenance, at least in principle, the “legal prohibition of ‘victimless’ immoralities.”³⁹ Reinterpreting the harm principle as a means-limiting doctrine of political legitimacy that ultimately serves the value of personal autonomy, Raz

³⁶ Raz, *Morality of Freedom*, 420.

³⁷ *Ibid.*, 417.

³⁸ *Ibid.*, 420.

³⁹ Robert P. George, *Making Men Moral* (Oxford: Oxford University Press, 1993), 167.

repudiates this illiberal approval of moral paternalism and firmly allies himself with liberalism.

III. DISCRIMINATION, AUTONOMY, AND HARM

To justify antidiscrimination law, Raz must confirm that the legislation satisfies the harm principle's legitimacy criterion, regulating behavior that actually causes harm. Moreover, because he conceives of the harm principle as an autonomy-promoting measure, Raz must ensure not only that discrimination is harmful, but also that antidiscrimination law promotes more autonomy than it curbs, that the amount of autonomy it secures for the victims of discrimination exceeds the amount of autonomy it denies discriminators. For Raz endorses the harm principle only on the assumption that it promotes autonomy, an assumption rooted in what John Stanton-Ife dubs the "asymmetry argument." According to that argument, "coercing to prevent harm leads to autonomy-gain, coercing to prevent (harmless) immorality leads to autonomy-loss; and that explains why legal coercion should be limited by harm-based considerations. It is for the sake of autonomy."⁴⁰ When this argument does not hold, the harm principle must defer to the value of autonomy, condemning the coercion as illegitimate—even if the coercion would prevent (perhaps considerable) harm. As Raz himself avers, we cannot sanction governmental coercion without considering both "the amount and nature of the harm to be prevented" and "the amount and nature of the harm coercion will inflict."⁴¹ Adopting an essentially consequentialist approach to political legitimacy, Raz thus incorporates into his version of the harm principle not only a necessary condition for legitimate governmental coercion, but also a proportionality requirement, a requirement antidiscrimination law must satisfy if it is to comport with his autonomy-based political morality.

⁴⁰ John Stanton-Ife, "The Limits of Law," *Stanford Encyclopedia of Philosophy*, <http://plato.stanford.edu/entries/law-limits/>. Cf. Raz, *Morality of Freedom*, 419.

⁴¹ Raz, *Morality of Freedom*, 421.

This proportionality requirement might seem to sit uncomfortably with Raz's broader moral theory, which unequivocally rejects consequentialism.⁴² And yet, Raz just as clearly countenances interpersonal tradeoffs of autonomy, maintaining that "[a] moral theory which values autonomy highly can justify restricting the autonomy of one person for the sake of the greater autonomy of others or even of that person himself in the future."⁴³ Raz can take this consequentialist approach to promoting autonomy because his account of rights entails no individual right to autonomy that would preclude sacrificing one person's autonomy for another's. More specifically, on Raz's account, a person has a right if and only if one of his interests "is a sufficient reason for holding some other person(s) to be under a duty."⁴⁴ Because a person's interest in being autonomous is never sufficient to justify burdening others with the demanding duty of helping to create an autonomy-promoting social environment, Raz argues, there can be no individual right to autonomy as such.⁴⁵ With no such right to constrain its pursuit of autonomy-based goals, the government may curb some citizens' autonomy so long as it offsets those restrictions with greater enhancements of other citizens' autonomy.⁴⁶ In the particular case of the legal prohibition of individual conduct, "Raz's account of rights does not provide any particular limit on criminalization, as the effectiveness of using criminal law to encourage people to act in accordance with their duties is always an empirical question. Empirically, it is entirely possible that extensive use of the criminal law might promote autonomy more than it interferes with autonomy."⁴⁷ When this condition is met, when a legal prohibition promotes more autonomy than it curbs, Raz's version of the harm principle

⁴² See *ibid.*, chs. 11-13.

⁴³ *Ibid.*, 419.

⁴⁴ *Ibid.*, 166.

⁴⁵ *Ibid.*, 247.

⁴⁶ See Hamish Stewart, "The Limits of the Harm Principle," *Criminal Law and Philosophy* 4 (2010): 25-26

⁴⁷ *Ibid.*, 26. Although Stewart uses the term "criminalization," Raz's main concern is coercion, which can be effected through means other than the criminal law.

will countenance the prohibition as a legitimate means of advancing his perfectionist political morality.

Departing from theories like Feinberg's, Raz's perfectionist political morality defines the concept of harm not by reference to just any of a person's interests, but rather exclusively in terms of the specific value of personal autonomy. To be sure, Raz initially echoes Feinberg's account of harm, claiming that "[t]o harm a person is to diminish his prospects, to affect adversely his possibilities. Roughly speaking, one harms another when one's action makes the other person worse off than he was, or is entitled to be, in a way which affects his future well-being."⁴⁸ But, Raz argues, unless we explicate the concept further in light of a particular theory of individual "well-being," of what it means for a person to be made "worse off," the harm principle will lack "concrete content," and we will be unable to apply it in particular cases.⁴⁹ Raz endows his version of the harm principle with "concrete content" by grounding his account of harm in his autonomy-based theory of well-being, equating harm with the impairment of autonomy. More specifically, one causes harm when one either erodes the preconditions of the autonomous life, "[d]epriving a person of opportunities or the ability to use them," or thwarts a person's pursuit of his autonomously chosen goals.⁵⁰ And because autonomy presupposes a hospitable social milieu, one in which everyone recognizes the value of autonomy and seeks to foster it, one can cause harm not just by directly interfering with a person's choices, but also by shirking one's more affirmative autonomy-based duties to help one's fellow human beings to lead autonomous lives—by, in other words, making people worse off than they are "entitled to be."⁵¹

⁴⁸ Raz, *Morality of Freedom*, 414.

⁴⁹ *Ibid.*

⁵⁰ *Ibid.*, 413.

⁵¹ *Ibid.*, 415-16.

Whereas some liberals might attempt to expand this account of harm to encompass certain other misfortunes that do not impair autonomy, Raz views only infringements of autonomy as harmful. In particular, Raz distinguishes between “the narrow harm principle which allows coercion only for the prevention of harm in the strict sense of the word [i.e., violations of autonomy] and the somewhat open-ended broad harm principle which allows coercion for the prevention of pain, offence, and perhaps some other injuries to a person as well.”⁵² Endorsing the former version of the harm principle, Raz refuses to countenance governmental coercion aimed at preventing harmless pain and offence, pain and offence that compromise no one’s autonomy.⁵³ Only if a given form of individual behavior impairs autonomy, Raz insists, may the government coercively curtail it.

Raz believes that discrimination based on race or sex does indeed impair autonomy and, consequently, that his version of the harm principle legitimates antidiscrimination law. For it seems that discrimination deprives its victims of opportunities other people readily enjoy and that considerably enhance the autonomy of their beneficiaries. Raz thus conceives of antidiscrimination law as but one facet of the broader “movement for social reform in the last two centuries” that “was motivated by a desire to make opportunities, which are available to some, open to all.”⁵⁴ More specifically, by “help[ing] to assure individuals that their opportunities will not be restricted by prejudiced denial of access to them,” the nondiscrimination norms that underwrite antidiscrimination law seek to promote the autonomy of particularly disadvantaged members of society, members who otherwise would lack the prerequisites for an autonomous life.⁵⁵ Moreover, Raz considers antidiscrimination law to epitomize

⁵² Ibid., 414.

⁵³ Ibid., 421.

⁵⁴ Raz, “Facing up,” 1229.

⁵⁵ Raz, “Liberty and Trust,” 115.

the kind of “legal intervention” that “can do some good,” that can advance the government’s autonomy-promoting agenda efficiently.⁵⁶ In Raz’s view, then, antidiscrimination law satisfies both the legitimacy criterion and the proportionality requirement embodied in his version of the harm principle.

Deprivation of Options

In fact, as John Gardner explains, discrimination based on race or sex seems to perpetrate a particularly serious violation of autonomy. For unlike other violations, discrimination conditions its victims’ access to various opportunities on their immutable traits, traits that are morally arbitrary and over which the victims themselves lack control. Discrimination thus at least partially subjects its victims to other people’s wills, preempting some of their most fundamental choices. This subordination, in turn, will compromise the victims’ autonomy when it deprives them of so many options in so many areas of their lives that their future paths are largely predetermined, that they can no longer shape their lives for themselves, but instead must simply acquiesce in others’ decisions about how they should act.⁵⁷

But will discrimination always be sufficiently severe for this harm to materialize? Do all instances of discrimination, in other words, necessarily deprive their victims of so many options that they can no longer lead autonomous lives? Although we might be tempted to evade these questions simply by deeming any deprivation of an option to be a violation of autonomy and thus a legitimate ground for legal coercion, Raz’s harm principle establishes a more stringent standard, recognizing only those deprivations that depreciate their victims’ choice sets below a certain sufficiency threshold. More specifically, his account of personal autonomy requires not that people have access to any

⁵⁶ Raz, “Facing up,” 1232.

⁵⁷ John Gardner, “On the Ground of Her Sex(uality),” *Oxford Journal of Legal Studies* 18 (1998): 170-71. See also John Gardner, “Private Activities and Personal Autonomy: At the Margins of Anti-Discrimination Law,” in *Discrimination*, eds. Bob Hepple and Erica M. Szyszczak (London: Mensell, 1992), 155.

particular opportunity, but that they enjoy merely an “adequate” range of valuable options.⁵⁸ In denying another person a particular opportunity, then, one will impair the other’s autonomy and thus harm him in the sense relevant for Raz’s harm principle only if one thereby renders his range of options inadequate. Similarly, if a discriminatory act extinguishes one option but leaves many others intact, it will likely not preclude his living an autonomous life and, according to Raz’s harm principle, will thus not merit a coercive legal remedy like antidiscrimination law.

This argument assumes a certain symmetry between Raz’s account of autonomy and his account of harm. In particular, it supposes that, if one needs only an adequate range of valuable options to be autonomous, then one’s autonomy cannot be compromised, one cannot be harmed, unless one’s choice set is somehow rendered inadequate. But perhaps such a symmetry between the two accounts is illusory. Perhaps Raz espouses a sufficiency standard in identifying the preconditions of autonomy, maintaining that the autonomous life requires only an adequate range of options, but then deems the denial of an option as such to be a violation of autonomy and thus harmful—even if the victim’s range of options remains technically adequate. Indeed, Raz himself seems explicitly to endorse such an asymmetry when articulating his definition of harm, noting that “[d]epriving a person of opportunities or the ability to use them is a way of causing him harm.”⁵⁹ It thus appears that Raz’s harm principle will consider harmful and subject to governmental regulation even those instances of discrimination that leave their victims with an adequate range of valuable options.

On closer inspection, however, we see that Raz’s broader theory cannot sustain such an asymmetry between his account of autonomy and his account of harm. For the former account holds that one does not necessarily impair another person’s autonomy

⁵⁸ See, e.g., Raz, *Morality of Freedom*, 372. Cf. Joseph Raz, *Ethics in the Public Domain*, rev. ed. (Oxford: Oxford University Press, 2001), 16-24.

⁵⁹ Raz, *Morality of Freedom*, 413.

simply by depriving him of a particular option,⁶⁰ while the latter maintains that one can harm another person only by impairing his autonomy (or by shirking one's duty to help him to become and to remain autonomous).⁶¹ Therefore, one impairs another person's autonomy and thereby harms him not whenever one deprives him of a particular option, but only when one renders his range of options inadequate, incapable of supporting an autonomous lifestyle. Only on this supposition can we explain Raz's conclusion that "[n]ot all coercion must infringe autonomy. Many cases of private coercion are very localized. They deny the coerced one option without interfering with his ability to choose from many others. Where those others are adequate no loss of autonomy is involved."⁶² In light of this specific conclusion, as well as his more general pronouncements, it becomes clear that Raz endorses a sufficiency criterion for both his account of autonomy and his account of harm.

It might be objected that the preceding argument rests on a false dichotomy, wrongly assuming that there are only two possible interpretations of Raz's account of harm: either every denial of a valuable opportunity is per se harmful or only those denials that render their victims' choice sets inadequate are. In fact, we might think that there are intermediate interpretations as well, ones that would deem some, but not all, denials of opportunities that left their victims with an adequate range of options to be harmful. For example, on Raz's autonomy-based theory, a person's well-being depends not just on the number and quality of the activities and relationships in which he participates, but also on the nature of his participation in those activities and relationships. In particular, activities and relationships will contribute to a person's well-being only if he *commits* to them, investing an appropriate amount of effort in pursuing them and appropriately privileging them over alternative valuable activities and relationships that he might have chosen to

⁶⁰ See, e.g., *ibid.*, 410, 425.

⁶¹ *Ibid.*, 416. The second kind of harm will be considered more closely below.

⁶² *Ibid.*, 421 n.1.

pursue instead.⁶³ If an option's contribution to a person's well-being thus depends on the nature of the person's engagement with the option, then so, too, might the harmfulness of the person's being deprived of the option. That is, a person might be harmed if he is deprived of a valuable option to which he has committed himself, even if he is left with an "adequate" range of other, equally valuable options that he had previously decided to forgo. As Raz puts it, "a harm to a person may consist not in depriving him of options but in frustrating his pursuit of the projects and relationships he has set upon."⁶⁴

This account, however, does not necessarily reject a sufficiency criterion for harm. Rather, it seeks to provide a more nuanced understanding of the threshold below which a choice set becomes inadequate. More specifically, it suggests that a person no longer enjoys an adequate range of valuable options if he is frustrated in the pursuit of an option to which he has committed himself, regardless of whether he continues to enjoy other options that are equally valuable but that he had previously forgone. We can assess the adequacy of a set of options both before and after a person has committed to any particular option in the set. Before a person has committed to any particular option in a set, the set's adequacy depends solely on the number and quality of the options it contains and on the person's ability to pursue them. Therefore, assuming that a set exceeds the threshold established by these criteria, any particular option can be eliminated without compromising the set's adequacy. By contrast, after a person has committed to a particular option in a set, the set remains adequate for him only so long as it contains the option to which he has committed; removing his chosen option renders the entire set inadequate, no matter how many other (perhaps equally) valuable options remain. For in committing to a particular option, a person inextricably links his well-being to his successful pursuit of the option, so that other options that were once equally eligible can

⁶³ See Raz, *Ethics in the Public Domain*, 5-8. See also Raz, *Morality of Freedom*, 381-90.

⁶⁴ Raz, *Morality of Freedom*, 413.

no longer serve as substitutes for the chosen option. With no meaningful substitutes for his chosen option, a person will enjoy an inadequate range of options if he is deprived of it. We can pronounce this kind of deprivation harmful while maintaining a threshold conception of harm.⁶⁵

Even if a threshold conception of harm accurately captures Raz's position as an interpretive matter, it still might be normatively objectionable. Perhaps the most serious objection to a threshold conception of harm is that it is underinclusive, that it condones deprivations of options that we intuitively believe are harmful. Suppose, for example, that a person enjoys a choice set that already lies below the adequacy threshold and then is deprived of an option.⁶⁶ Or consider a deprivation suffered by a person whose choice set so exceeds the adequacy threshold that even a serious deprivation will not render the choice set inadequate.⁶⁷ On Raz's account of harm, neither of these deprivations is harmful, since neither depreciates the victim's choice set below the adequacy threshold. And yet, at least in certain circumstances, both deprivations would strike us as harmful.

One possible response to this objection is to concede that Raz's threshold conception of harm does indeed fail to condemn certain objectionable deprivations of options as harmful yet to insist that such a result is less damning than it initially seems. After all, in failing to deem a given deprivation harmful, Raz's version of the harm principle precludes not all governmental responses to the deprivation, but rather only those responses that involve coercion. For example, although the government could not legally prohibit conduct that merely deprived people of options without rendering their

⁶⁵ The argument of this paragraph assumes that the option to which the person has committed is a significant one—say, a particular career choice. Of course, if the option to which a person commits is only a trivial one, then removing it from the person's choice set does not necessarily render the set inadequate.

⁶⁶ Cf. Lukas H. Meyer, "Past and Future: The Case for a Threshold Notion of Harm," in *Rights, Culture, and the Law*, eds. Lukas H. Meyer, Stanley L. Paulson, and Thomas W. Pogge (Oxford: Oxford University Press, 2003), 153-54. Note that Meyer's threshold conception of harm is concerned with levels of overall well-being, whereas Raz's is concerned with autonomy.

⁶⁷ Cf. *ibid.*, 154.

choice sets inadequate, it could provide the victims of such conduct with subsidies in order to restore their choice sets to the status quo ante. Given the availability of such alternative forms of governmental regulation, the failure of Raz's threshold conception of harm to condemn certain objectionable deprivations of options as harmful, and thus to authorize the coercive regulation of such deprivations, need not prompt us to jettison that conception for a more expansive one.

Raz's version of the harm principle might even be able to authorize the government to use coercion to remedy some deprivations of options that do not depreciate their victims' choice sets below the adequacy threshold. For according to Raz, one can cause harm not just by depriving others of options, but also by shirking one's duties to help to make others autonomous.⁶⁸ To be sure, as we saw above, Raz rejects any general duty to help to make others autonomous, considering such a duty to be too demanding. But we might nevertheless have other, more specific and less demanding autonomy-based duties, such as a duty to minimize the injury we cause others by depriving them of options.⁶⁹ Such a duty might require us to avoid depriving another person of an option, even where the deprivation would leave the other person with an adequate range of options. And because the violation of such a duty would, by definition, be harmful on Raz's theory, the government could legitimately use coercion to prevent it.⁷⁰ Assuming Raz's theory can justify such a duty, the failure of his threshold account of harm to condemn certain objectionable deprivations of options as harmful is unproblematic. Nevertheless, it remains the case that, on Raz's account, a deprivation of an option will cause harm *directly*, irrespective of any duty to minimize the injury one causes, only if it renders the victim's choice set inadequate.

⁶⁸ See Raz, *Morality of Freedom*, 415-16, 417.

⁶⁹ See Meyer, "Past and Future," 153-54.

⁷⁰ A version of this argument will be considered below. For now, note that its success will depend on the reasonableness of imposing on everyone a duty always to avoid depriving others of options, regardless of the deprivation's severity.

To determine whether particular instances of discrimination actually cause such harm, we must first identify more precisely the threshold a choice set must attain to qualify as “adequate.” According to Raz’s theory of autonomy, there are a number of respects in which a person’s choice set can fail to be adequate. First, a person’s choice set will be inadequate if it includes only “trivial” options, options that are “short-term and negligible in their significance and effects.”⁷¹ For unless a person enjoys “options with long term pervasive consequences as well as short term options of little consequence, and a fair spread in between,” he will lack control over certain areas of his life and, consequently, full autonomy.⁷² Second, a person’s autonomy will be similarly diminished if virtually all his options are “potentially horrendous in their consequences”—if, for example, only one of his options affords him a significant chance of preserving his life.⁷³ Third, a person’s options should be sufficiently diverse that he has the opportunity to cultivate and to exercise his various capacities.⁷⁴ Finally, as we observed above, once a person commits to a particular option (or at least a significant one), his choice set will be adequate only if it contains the chosen option.⁷⁵ Note, however, that none of these criteria imply that a choice set must contain numerous options to be adequate. On the contrary, we frequently prefer relatively limited choice sets to more extensive ones—a choice between two different works of fiction, say, to one among fifty virtually indistinguishable editions of the same title.⁷⁶ To be sure, a person must confront more than just two alternatives to be autonomous, but autonomy can nonetheless subsist on relatively few options.

⁷¹ Raz, *Morality of Freedom*, 374.

⁷² *Ibid.*

⁷³ *Ibid.*, 374, 376.

⁷⁴ *Ibid.*, 375.

⁷⁵ *Ibid.*, 376.

⁷⁶ Cf. Philip Pettit, “Agency-Freedom and Option-Freedom,” *Journal of Theoretical Politics* 15 (2003): 392.

Given these criteria, it seems that many instances of discrimination will not depreciate their victims' ranges of options below the adequacy threshold and, consequently, that Raz's version of the harm principle will forestall many applications of antidiscrimination law. Indeed, Raz's version of the harm principle succumbs to precisely the same objection that undermined Feinberg's in the previous chapter, for in espousing a sufficiency standard of harm, it likewise overlooks discriminatory acts that affect their victims' well-being only modestly. Just like Feinberg, Raz refuses to deem the mere denial of an option harmful, insisting instead that "[t]he significance of denial of options to one's autonomy depends on the circumstances one finds oneself in."⁷⁷ Therefore, only if a denial renders the victim's range of options inadequate, hindering the victim's ability to lead an autonomous life, will Raz's version of the harm principle pronounce the denial harmful and license the government to remedy it through legal coercion.

Raz himself appears to concede that at least some instances of discrimination will not satisfy this requirement, thereby implying that at least some applications of antidiscrimination law are illegitimate. According to Raz, discrimination based on race or sex places "extraneous" limitations on access to opportunities—limitations "external to the activity concerned" rather than limitations "based on relevant qualifications."⁷⁸ However repugnant these limitations might be, Raz argues that they do not necessarily curb autonomy, since they "can leave everyone with more than adequate access" to other options.⁷⁹ When discriminatory limitations have only this more modest impact, they will not cause harm and thus will not warrant the applications of antidiscrimination law designed to eradicate them.

We have good reason to suppose that discrimination will indeed leave its victims with adequate ranges of options in many circumstances. For as we discovered in the

⁷⁷ Raz, *Morality of Freedom*, 410.

⁷⁸ Raz, *Ethics in the Public Domain*, 22.

⁷⁹ *Ibid.*

previous chapter, we can envision a society in which many members of previously oppressed social groups have attained levels of well-being that render them less vulnerable to the mere deprivation of an option, less dependent on any one opportunity in particular. Such a society might also have succeeded in largely marginalizing discrimination. To be sure, perhaps not all racial minorities and women in such a society will have gained adequate access to the opportunities most of their fellow citizens enjoy, and perhaps not all communities in such a society will have combated racism and sexism with the requisite resolve. Where racial minorities and women enjoy more abundant opportunities and ebbing racism and sexism, however, discrimination will be much less likely to deprive its victims of adequate ranges of valuable options and, consequently, much less likely to diminish their autonomy. Raz's version of the harm principle will condemn many applications of antidiscrimination law in such circumstances.

For at least two reasons, Raz's version of the harm principle will entail this result even though discrimination might sometimes frustrate its victims' pursuit of options to which they have committed themselves. First, we might question the conclusion, reached above, that a person is necessarily harmed whenever he is frustrated in his pursuit of an option to which he has committed himself, irrespective of the other options he continues to enjoy. This conclusion is plausible where the thwarted opportunity is a person's career, but it is less plausible where the thwarted opportunity is one of multiple hobbies to which a person has committed himself equally. That is, the extent to which a person's well-being depends on the success of his chosen projects itself depends on the significance of those projects.⁸⁰ Nevertheless, it does seem that frustrating a person in his pursuit of an option to which he has committed himself will often render his choice set inadequate and therefore harm him.

⁸⁰ Cf. Raz, *Morality of Freedom*, 376 ("Since people's well-being depends to a considerable extent on the projects and relationships *around which their lives revolve*, frustrating their successful pursuit undermines people's well-being." (emphasis added)).

But second, and more significant, discrimination does not always frustrate its victims' pursuit of options to which they have committed themselves. For example, in seeking employment, some people commit themselves to a particular job opportunity, expending all their effort on securing it and forgoing any comparable alternatives. But many other people make no such commitment, applying instead for a range of roughly equivalent jobs and forming no particular preference among them. In this latter case, discrimination will not harm its victims, since, *ex hypothesi*, they never committed themselves to the denied employment opportunity and continue to enjoy a significant number of equally desirable alternatives. By contrast, firing a person for a discriminatory reason will harm him, since he will have necessarily committed himself to his job by accepting and remaining in it. Raz's harm principle thus entails an asymmetric approach to discriminatory hiring and firing, frequently condoning the former conduct but generally condemning the latter. It will not, however, condemn those (admittedly rare) discriminatory firings where the terminated employee is no longer committed to his job and thus no longer views it as one of his life projects—say, where he plans to resign anyway and has already secured new employment. Nor will it condemn most discriminatory exclusions from public accommodations, to which people tend not to commit themselves in the way they commit to their jobs. In short, because we do not always commit ourselves to the options we enjoy in the contexts governed by antidiscrimination law, the victims of discrimination will not always be harmed when they are deprived of an opportunity in those contexts.⁸¹

Like an antiperfectionist version of the harm principle, then, Raz's perfectionist one can countenance not the continuous antidiscrimination regimes ascendant in most

⁸¹ This paragraph has ignored instances of discrimination that deprive their victims of *valueless* options. Obviously, because valueless options do not contribute to autonomy, their denial is not harmful, according to Raz's theory. For example, assuming employment in the pornography industry to be valueless, the discriminatory denial of a role in a pornographic film does not harm its victim.

western liberal democracies today, but only more patchwork regulation, regulation that applies only in certain communities and that protects only certain segments of the population. And yet, Raz's version of the harm principle does not constrain other facets of the government's nondiscrimination agenda nearly so severely as would an antiperfectionist one. For whereas the latter would categorically prohibit the government from attempting to curb harmless instances of discrimination, the former merely forbids it to use coercion for that purpose. In other words, even when discrimination leaves its victims with an adequate range of options, impairing no one's autonomy and thus causing no harm, a government limited by Raz's harm principle may nonetheless seek to dissuade the discriminators from their wrongdoing by remonstrating with them, publicly shaming them, taxing them, or subsidizing their economic competitors—in short, by employing any means short of coercion.⁸² Raz's perfectionist version of the harm principle thus affords the government considerable latitude in addressing harmless instances of discrimination. But even this flexibility cannot accommodate all applications of antidiscrimination law, some of which employ coercion not to promote autonomy, but merely to censure immoral conduct—an exercise of coercion that even Raz's harm principle expressly forbids.

Compromised Self-respect and Second-class Citizenship

The argument thus far has considered only how discrimination might restrict its victims' opportunities, ignoring the conduct's impact on the victims themselves. As noted above, however, one can violate another person's autonomy not only by depriving him of an adequate range of valuable options, but also by eroding his personal capacity to make autonomous choices. Raz believes that discrimination tends to compromise autonomy in the second way as well as the first. More specifically, by impugning the identities of the

⁸² A critic of Raz's theory might question whether these alternative measures really harm autonomy any less than coercion does. See, e.g., Jeremy Waldron, "Autonomy and Perfectionism in Raz's *Morality of Freedom*," *Southern California Law Review* 62 (1989): 1141-52.

social groups to which its victims belong, discrimination renders the groups objects of ridicule and scorn rather than of admiration and respect, preventing many of their members—especially those who have suffered discrimination themselves, but also some who have not—from “feel[ing] pride” in their membership.⁸³ This sense of alienation, in turn, tarnishes the members’ own personal identities, since most people consider identification with various social groups to be “an important element in their life.”⁸⁴

With their identities thus insulted, the members of the derided social groups cannot enjoy complete autonomy. For the autonomous life assumes a certain “emotional and imaginative make-up” in its subjects, a certain self-conception compatible with the ideal of autonomous agency.⁸⁵ In particular, one cannot be autonomous unless one has “self-respect,” “accept[ing] without alienation one’s core being, one’s core pursuits and relationships and those aspects of one’s character and circumstances that one identifies with most deeply.”⁸⁶ If, by contrast, one comes to view oneself with a sense of “shame” or “self-hate,” one will be unable to approach one’s various choices, projects, and relationships with the mindset autonomy requires. Discrimination based on race or sex engenders precisely these pathological feelings in many racial minorities and women by implying the inferiority of the social groups to which they belong, regardless of whether it also restricts anyone’s range of options.⁸⁷

Moreover, by branding certain social groups and their members as inferior, discrimination also adulterates society’s broader public culture, distorting the terms on which citizens relate to one another. More specifically, many members of disfavored social groups will come to consider themselves, and to be perceived by their compatriots

⁸³ Raz, *Morality of Freedom*, 254.

⁸⁴ *Ibid.*

⁸⁵ *Ibid.*, 408.

⁸⁶ Raz, *Ethics in the Public Domain*, 25, 26.

⁸⁷ *Ibid.*, 25.

as, second-class citizens.⁸⁸ Such subordination, in turn, will further diminish autonomy, which presupposes one's "feeling part of a larger community, and being able to identify oneself as a member of such communities."⁸⁹ Were it to ignore these deleterious effects of discrimination and to decline to enact antidiscrimination legislation to counteract them, the government would display a general disregard for the well-being of racial minorities and women, only exacerbating those persons' sense of alienation from their society. Indeed, rather than viewing their political community's "law as their law, and its government as their government," many of them would likely come to regard the institutions of the state as nothing more than "sectarian" instruments through which the members of other, more privileged social groups pursue their parochial ends.⁹⁰ Because such beliefs undermine equal citizenship and thus preclude full autonomy, Raz's harm principle authorizes the government to combat them.

Although Raz's harm principle thus authorizes the government to proscribe objectionable conduct that imperils people's self-respect and equal citizenship, we might nevertheless question whether discrimination always has these profound psychological effects—whether, in particular, it always conveys so potently denigrating a message that it prevents not just its instant victims, but whole classes of citizens, from taking pride in important aspects of their identities and thus from enjoying full autonomy. In fact, to attribute such meaning to all discrimination is to exaggerate the conduct's expressive power. To be sure, even in a reasonably tolerant society, discrimination will continue to affront some of its victims and—more obliquely—other people who share their racial or gender identities. But it nonetheless seems unlikely that this insult will be sufficiently conspicuous and cutting to diminish many people's autonomy. For many discriminators in such a society will be motivated not by naked racism and sexism, but by economic

⁸⁸ Raz, "Liberty and Trust," 126.

⁸⁹ *Ibid.*, 125.

⁹⁰ *Ibid.*, 124.

considerations or even sheer ignorance,⁹¹ while most of the remaining incorrigible racists and sexists will attempt to dissemble their wrongdoing, thereby occluding any malicious message they might otherwise impart. Indeed, Raz himself suggests that racist and sexist “attitudes” must permeate the broader public culture and that whole “societies,” not just individual discriminators, must espouse racism and sexism if racial minorities and women are to suffer significant harm to their self-respect, to their sense of equal citizenship, and, hence, to their autonomy from discrimination.⁹²

But even if discrimination dejected some racial minorities and women so deeply that they lost self-respect and thus became less than fully autonomous, it does not follow that Raz’s harm principle would legitimate all applications of antidiscrimination law. For other racial minorities and women might not react so severely to discriminatory conduct—whether perpetrated against them or against other members of their social groups. Recall that, on Raz’s account, it is not enough that racial minorities and women be disappointed, upset, embarrassed, angry, or ashamed by discrimination. Rather, they must experience particularly intense emotions (Raz uses terms like “self-hate” and “alienation”), emotions sufficiently intense to compromise one’s ability to lead an autonomous life, in order to be harmed in the sense relevant for the harm principle. Whether racial minorities and women will experience these especially intense, autonomy-compromising emotions or only more modest ones will depend on at least two factors. First, as with any group of human beings, some racial minorities and women will, for whatever reason, be less emotionally vulnerable than others. Second, it seems that well-off racial minorities and women will generally suffer less psychological injury from discrimination than their more disadvantaged counterparts, since enjoying abundant options and multiple bases of self-respect can significantly diminish (though, admittedly,

⁹¹ We should not, of course, infer either that the discriminators are any less blameworthy or that their wrongdoing is any less immoral on that count.

⁹² See Raz, *Ethics in the Public Domain*, 26.

never eliminate) the insult that discrimination conveys. Because Raz's autonomy-based account of harm will not consider such instances of psychological injury to be harmful, his version of the harm principle will fail to legitimize some applications of antidiscrimination law.

Just how many applications of antidiscrimination law Raz's version of the harm principle will forestall is a difficult empirical question whose answer will depend on the socioeconomic composition of minority groups. If most racial minorities, for instance, suffer from socioeconomic disadvantage, then Raz's version of the harm principle will be able to legitimize nearly all applications of antidiscrimination law using a version of the types-tokens strategy considered in the previous chapter, since discrimination will be sufficiently likely to cause harm to justify prohibiting all its instantiations. But as the proportion of racial minorities who are well-off increases, the types-tokens strategy will become increasingly implausible, and Raz's version of the harm principle will legitimize ever fewer applications of antidiscrimination law. This result owes to Raz's consequentialist approach to political legitimacy: only when the vast majority of racial minorities and women are disadvantaged, so that the psychological injury wrought by discrimination is likely to compromise their autonomy, will Raz's version of the harm principle legitimate all applications of antidiscrimination law. Once again, it becomes clear that, at least in certain social conditions, Raz's harm principle will justify at best a sporadic antidiscrimination regime.

Antidiscrimination Law: Political as Well as Moral?

We might think that the preceding criticisms unrealistically expect Raz's theory to reconcile considerations of moral principle with those of political prudence when, in fact, the two kinds of considerations inevitably conflict. More specifically, in insisting that Raz's autonomy-based political morality furnish purely principled, rather than even partly

pragmatic, justifications for *all* applications of antidiscrimination law, applications both to privileged racial minorities and women as well as to disadvantaged ones, the foregoing arguments might appear to ignore the considerable institutional and other prudential imperatives that necessarily constrain the implementation of any moral agenda. Raz thus might readily concede that his policy program outstrips the justificatory force of his moral theory, extending the protections of antidiscrimination law to persons who are technically not entitled to them on grounds of autonomy. But why, he might ask, should this incongruity in any way impugn his autonomy-based harm principle? If we cannot guarantee autonomy to some persons without simultaneously benefiting others who already possess it, if, for example, antidiscrimination law cannot secure for certain disadvantaged persons an adequate range of options or a robust sense of self-respect without simultaneously protecting other, privileged persons who already enjoy these goods, then it seems utterly obtuse for the harm principle to condemn the non-harm-preventing applications of antidiscrimination law on which the efficacy of harm-preventing ones depends. We have good reason to suppose, moreover, that harm-preventing and non-harm-preventing applications of antidiscrimination law are closely allied, for it seems rather unlikely that public opinion would ever support a bifurcated antidiscrimination regime, one that protected only disadvantaged racial minorities and women while spurning privileged ones, or that the law could even draw the fine distinctions that such a regime would require.⁹³ Given these inherent limitations of politics and the law, Raz can plausibly argue that, despite any qualms it might have about particular applications of antidiscrimination law, his version of the harm principle must honor the obligation to secure the autonomy of especially disadvantaged racial minorities

⁹³ Cf. *ibid.*, 172-73 (adducing certain political and other pragmatic considerations to defend a policy of toleration, which Raz then depicts as the precursor of nondiscrimination).

and women, an obligation that can be discharged only through a uniform, continuous antidiscrimination regime.

This line of reasoning, which resembles the kind of two-level logic considered in Chapter 1, bespeaks a more fundamental objection to the orthodox understanding of antidiscrimination law. Whereas we traditionally view antidiscrimination law as deeply principled legislation, legislation that seeks to transpose fundamental moral values like autonomy directly into public policy, the argument expounded in the preceding paragraph portrays at least some central facets of antidiscrimination regimes as little more than concessions to political and prudential realities. To be sure, even these more prudential aspects still advance pressing moral goals, since they not only facilitate the deeply principled applications of antidiscrimination law, but also thwart discriminatory acts that, while perhaps harmless to autonomy, contravene other fundamental values. But at least within Raz's moral theory, they appear to serve a largely ancillary function, responding to some of the institutional and political challenges that threaten to undermine antidiscrimination law's principal, autonomy-promoting mission. An effective antidiscrimination regime incorporates both principled and prudential components, and we must recognize both if we are comprehensively to justify antidiscrimination law.

At first glance, Raz's perfectionist political morality seems particularly well-equipped to support this two-level justification of antidiscrimination law. For, as we saw earlier, Raz believes that perfectionism affords the government considerable latitude in pursuing its regulatory agenda, pronouncing "all moral reasons . . . fair game for governmental action."⁹⁴ Unlike an antiperfectionist government, then, a perfectionist one may invoke not just some circumscribed set of reasons, but all sound reasons when justifying its policies and thus may shape the implementation of its goals according to

⁹⁴ Raz, "Facing up," 1230.

various considerations an antiperfectionist government must disregard. In particular, because it conceives of doctrines of political legitimacy not as categorical injunctions against certain governmental actions, but rather as prudential guidelines for realizing particular moral goals, perfectionism permits the government occasionally to relax the harm principle to accommodate pragmatic imperatives, flexibility that antiperfectionism precludes. It thus appears that Raz can justify the full extent of contemporary antidiscrimination law so long as he can adduce a harm-preventing rationale for at least the legislation's core applications—even if its more peripheral applications rest primarily on prudential considerations.

On closer inspection, however, Raz's version of the harm principle turns out to be much less amenable to the kind of two-level logic that this defense of antidiscrimination law employs. As noted earlier, Raz's version of the harm principle functions as a means-limiting doctrine of political legitimacy, restricting the means the government may use to further its adopted ends rather than specifying the ends it may adopt in the first place. In particular, although Raz's harm principle permits the government to pursue any (morally sound) end, even a perfectionist one, it permits the government to use coercion only to prevent harm, understood as an impairment of autonomy. And yet, this classification of Raz's version of the harm principle as a means-limiting doctrine of political legitimacy is a little too neat, for while the doctrine does not limit the reasons the government may adduce for its policies in general, it does limit the reasons the government may adduce for its use of coercion in particular. Raz's harm principle thus functions as an ends-limiting doctrine of political legitimacy with respect to the government's exercise of coercion, if not with respect to its more general policy program.

Insofar as it functions as an ends-limiting doctrine of political legitimacy (even if only with respect to the government's use of coercion), Raz's version of the harm

principle has difficulty adopting the two-level logic that a pragmatic account for antidiscrimination law seems to presuppose. Recall that, according to that account, the government may legitimately prohibit harmless instances of discrimination because people would not support a bifurcated antidiscrimination regime, one that prohibited only genuinely harmful instances of discrimination and that consequently protected only disadvantaged racial minorities and women. The rationale for antidiscrimination law would, on this view, remain an autonomy-based one, since only by prohibiting instances of discrimination that did not compromise their victims' autonomy could the government effectively combat those that did.

But why would people oppose a bifurcated antidiscrimination regime if, in fact, discrimination harmed only particular subpopulations of racial minorities and women? Two reasons seem most plausible. First, especially in a highly class-conscious society, the well-off members of all races and both sexes might resent the government's extending antidiscrimination protections only to the disadvantaged, believing that such protections should not be apportioned on a socioeconomic basis. But as should now be clear, such a belief would be unsound according to Raz's theory of autonomy, which holds that people who are disadvantaged—who lack the preconditions for an autonomous life—may legitimately demand from the government benefits to which their more well-off compatriots are not entitled. To adopt a two-level justification of antidiscrimination law simply in order to accommodate this erroneous belief would be to compromise a morally sound theory for the sake of moral error, producing the kind of incoherence attributed to two-level doctrines of political legitimacy in Chapter 1.

Second, people might oppose a bifurcated antidiscrimination regime not because they resent missing out on benefits enjoyed by others, but because they believe that discrimination is always wrong, irrespective of the effects it has on its victims. Of course,

if this belief is erroneous, if discrimination is wrong only when it compromises its victims' autonomy, then to adopt a two-level justification of antidiscrimination law in order to accommodate it would again be to render Raz's theory incoherent. But the threat of incoherence remains even if discrimination is indeed always wrong, irrespective of its effects. For Raz's harm principle requires the government to adduce an autonomy-based justification for the use of coercion. To comply with this requirement while implementing an antidiscrimination regime that was not, strictly speaking, necessary to promote autonomy, the government would have to adopt something like the following line of reasoning: "Only a bifurcated antidiscrimination regime is necessary to promote autonomy, but people (rightly) believe that discrimination is always wrong, irrespective of its effects, and will consequently support only a uniform, continuous antidiscrimination regime. Thus, we can best promote autonomy only by acquiescing in people's morally sound, but nevertheless non-autonomy-based, beliefs." The problem with this line of reasoning is that it holds the integrity of Raz's harm principle hostage not to erroneous beliefs, but to beliefs that, on Raz's theory, are irrelevant to assessing the legitimacy of the government's use of coercion. If Raz's harm principle rightly deems the immorality of a given form of conduct to be irrelevant to determining whether the conduct may be prohibited, then it should discount people's beliefs about the conduct's immorality, not uncritically acquiesce in them.⁹⁵

This conclusion highlights an important affinity between Raz's perfectionist version of the harm principle and an antiperfectionist version like the one considered in

⁹⁵ Raz could continue to insist on catering to people's irrelevant and erroneous beliefs only by calling his theory's liberalism into question. To be sure, unlike an antiperfectionist version of the harm principle, Raz's version can distinguish between people's legitimate and illegitimate beliefs. But if his version of the harm principle is in principle willing to bend to accommodate people's beliefs, whatever their content, then it risks countenancing governmental regulation that nearly all liberals condemn. Suppose, for instance, that, in light of people's beliefs, the government could best promote autonomy by prohibiting same-sex sexual conduct. Raz's version of the harm principle would appear to legitimize the prohibition in such circumstances, a result that violates the spirit of the harm principle and liberalism more generally. Cf. Stewart, "Limits," 28-30.

the previous chapter. Although the former version generally permits the government to consider reasons the latter one suppresses, they both seem to entail remarkably similar regimes of coercion. For, by espousing the harm principle and then defining “harm” narrowly in terms of autonomy, Raz’s theory restricts the government’s use of coercion almost as severely as antiperfectionism does. To be sure, unlike an antiperfectionist version of the harm principle, Raz’s perfectionist version still permits the government to invoke non-autonomy-based reasons to justify its noncoercive policies. But both antiperfectionism and Raz’s theory render these other reasons irrelevant to the justification for coercion. Only by permitting the government to justify coercion according to values other than autonomy could Raz further distinguish his theory from antiperfectionism and thereby legitimate more applications of antidiscrimination law.

Shirked Duties

Before abandoning Raz to the same fate that befell antiperfectionists in the previous chapter, however, we must recall that his account of harm also condemns some forms of conduct that do not, strictly speaking, impair autonomy. His version of the harm principle thus might legitimate significantly more applications of antidiscrimination law than the relatively conventional account of harm considered above seemed to suggest it could. As we discovered earlier, Raz believes that one can harm another person not just through one’s acts, but also through one’s omissions—not just by impairing the other’s autonomy, but also by failing to enhance it when autonomy-based duties so require.⁹⁶ This symmetry between acts and omissions reflects a moralized conception of harm according to which one harms another person whenever one shirks the autonomy-based duties one owes him—whether one thereby diminishes his preexisting level of autonomy

⁹⁶ See Raz, *Morality of Freedom*, 415-16, 417.

or hinders his attaining a higher level of autonomy to which he is “entitled.”⁹⁷ Because Raz thus pronounces all violations of autonomy-based duties harmful, his version of the harm principle licenses the government “to use coercion both in order to stop people from actions which would diminish people’s autonomy and in order to force them to take actions which are required to improve peoples’ [*sic*] options and opportunities.”⁹⁸

When considered in light of the government’s imperative to enforce its citizens’ more positive autonomy-based duties, antidiscrimination law assumes a new objective: compelling people to properly respect the autonomy of racial minorities and women. More specifically, antidiscrimination law seeks to ensure that employers, purveyors of public accommodations, and other relevantly situated persons provide valuable opportunities to racial minorities and women, helping them to secure the adequate range of options autonomy requires. While prohibitions against direct discrimination appear to advance this aim, it is prohibitions against indirect discrimination that most evidently embody our affirmative autonomy-based duties. Recall that, according to U.K. antidiscrimination law, one indirectly discriminates against another person based on race or sex if one “applies” to her a “provision, criterion or practice” that one would apply equally to people of other races or the other sex but that nevertheless disproportionately disadvantages members of her race or sex and for which one can offer no compelling race- or sex-neutral justification.⁹⁹ So defined, indirect discrimination violates autonomy-based duties, for in promulgating, say, an ostensibly race-neutral hiring requirement that nevertheless disproportionately disadvantages racial minorities, an employer displays insufficient concern for the autonomy of the members of the previously oppressed groups. Prohibitions against indirect discrimination attempt to counteract such indifference by enjoining relevantly situated persons to facilitate racial minorities’ and women’s access to

⁹⁷ Ibid., 414.

⁹⁸ Ibid., 416.

⁹⁹ Equality Act 2010, § 19.

various opportunities.¹⁰⁰ In thus enhancing people's autonomy, these prohibitions also enlist their subjects in "a participative enterprise of protecting autonomy, an enterprise which carries with it obligations of mutual life-enhancement."¹⁰¹ Prohibitions against indirect discrimination (as well as those against direct discrimination) thus further a perfectionist mission, prodding people to respect and promote the autonomy that makes others' lives, as well as their own, so valuable.

Before approving this mission, however, we must determine how onerous the autonomy-based duties imposed by antidiscrimination actually are—in particular, whether they are so burdensome that they compromise their bearers' own autonomy by precluding the devotion to personal projects and relationships that the autonomous life requires. For, as we discovered earlier, Raz's harm principle requires not only that governmental coercion prevent harm, but that it do so economically, that it not cause even more harm in the process. The legitimacy of governmental coercion, in other words, depends on both "the amount and nature of the harm to be prevented" and "the amount and nature of the harm coercion will inflict."¹⁰² Therefore, if coercively enforcing autonomy-based duties harms the bearers of the duties more than it helps their intended beneficiaries, if it entails losses for some people's autonomy without achieving offsetting gains for that of others, then Raz's version of the harm principle will pronounce the coercion illegitimate. Indeed, according to Raz's autonomy-based morality, duties that prove too onerous are not really duties at all, since, Raz argues, "[w]hether or not . . . a duty is justified depends not only on the interest of the right-holder but, at the very least, also on that of the person who is to be subject to that duty."¹⁰³ Similarly, duties of

¹⁰⁰ Cf. Gardner, "On the Ground," 182-83.

¹⁰¹ John Gardner, "Liberals and Unlawful Discrimination," *Oxford Journal of Legal Studies* 9 (1989): 19.

¹⁰² Raz, *Morality of Freedom*, 421.

¹⁰³ Raz, *Ethics in the Public Domain*, 35.

nondiscrimination will prove unjustified if one cannot comply with them without excessively restricting one's own autonomy.

We might worry that, in evincing such solicitude for the autonomy of discriminators, we seriously misapprehend Raz's autonomy-based political morality. For, as Raz repeatedly avers, his commitment to autonomy is an unequivocally perfectionist one, according to which "[a]utonomy is valuable only if exercised in pursuit of the good."¹⁰⁴ The imperative to respect and promote autonomy, in other words, "does not extend to the morally bad and repugnant. Since autonomy is valuable only if it is directed at the good it supplies no reason to provide, nor any reason to protect, worthless let alone bad options."¹⁰⁵ On this view, one diminishes another person's autonomy not whenever one denies the other person an option, but rather only when the denied option is a valuable one. Therefore, when the government must coercively deprive some people of valueless or corrupt options in order to prevent harm to others, it need not ensure that the latter, autonomy-promoting consequences outweigh the former, autonomy-restricting ones in order to justify the coercion. Nor need it perform such a calculus when, in coercively compelling people to discharge their affirmative autonomy-based duties, it impedes their pursuit of valueless or corrupt options. It likewise seems that the government need have no compunction about coercively compelling discriminatory employers and purveyors of public accommodations to honor their autonomy-based duties, coercion that apparently forecloses only the corrupt option of discriminating against racial minorities and women.

Although this line of reasoning rightly emphasizes the perfectionist limits of Raz's commitment to autonomy, it nevertheless reflects a facile view of the relationships and activities in which discriminatory employers and purveyors of public accommodations

¹⁰⁴ Raz, *Morality of Freedom*, 381.

¹⁰⁵ *Ibid.*, 411.

often are engaged. More specifically, in assuming that prohibitions against discrimination curtail only valueless or corrupt options, it disparages the worth and complexity of those relationships and activities and thus downplays the often-considerable costs antidiscrimination law has for autonomy, understood in purely perfectionist terms. After all, discriminators seek to exclude racial minorities and women not necessarily to harm them gratuitously (though some discriminators undoubtedly do have such callous motives), but often to shape some of their own concededly valuable relationships and activities according to their own personal preferences. If antidiscrimination law cannot secure the autonomy of racial minorities and women without subverting these relationships and activities, it will necessarily restrict not only vicious options, but also valuable ones.¹⁰⁶ This harm will be particularly acute, moreover, if antidiscrimination law imposes autonomy-based duties so onerous that their bearers must forgo certain valuable options in order to discharge them. We thus cannot simply assume that antidiscrimination law augments the valuable autonomy of the victims of discrimination without curtailing that of the perpetrators. Rather, we must weigh the legislation's effects on the former against its effects on the latter, a calculus that even Raz's perfectionist harm principle does not obviate.

The results of this calculus will depend largely on how Raz's theory specifies our autonomy-based duties—in particular, those autonomy-based duties imposed by antidiscrimination law. It seems that Raz's theory is ambiguous between two understandings of those duties, one of which renders the duties considerably less burdensome than the other. First, autonomy-based duties might require employers and purveyors of public accommodations to provide opportunities to the racial minorities and women they encounter only when the latter persons lack an adequate range of options,

¹⁰⁶ Cf. *ibid.*, 418-19.

only, that is, when the latter persons enjoy so few opportunities that they cannot be fully autonomous. When, by contrast, a racial minority or a woman already possesses an adequate range of options, the employers to whom she applies and the purveyors of public accommodations whose services she solicits have no autonomy-based duty to offer her the opportunity she seeks. According to this interpretation, then, one violates no autonomy based duty and thus causes no harm simply by withholding a particular opportunity from another person unless one thereby impedes the other person from securing an adequate range of options.

If Raz does indeed understand autonomy-based duties in this way, then his theory will render many applications of antidiscrimination law illegitimate. For, as we saw above, when racial minorities and women possess adequate ranges of options, they will not be harmed by most discriminatory acts. By enjoining employers and purveyors of public accommodations to provide opportunities only to those racial minorities and women who lack adequate ranges of options, autonomy-based duties would impose obligations significantly less stringent than the ones imposed by modern antidiscrimination regimes, which proscribe discrimination against even the most affluent racial minorities and women. It thus becomes clear that Raz cannot espouse a more modest conception of autonomy-based duties without foundering on the very same empirical reality that scuttled his more conventional account of harm above—namely, the fact that some discriminatory acts will simply not impair their victims' autonomy and thus will cause no harm.

To vindicate more applications of antidiscrimination law, we alternatively might construe the autonomy-based duties imposed by the legislation to be significantly more onerous. The duties might, in particular, require employers and purveyors of public accommodations to provide opportunities not merely to those racial minorities and

women who lack adequate ranges of valuable options, but to all members of the historically disadvantaged groups—regardless of the members’ individual levels of well-being. So understood, the duties would indeed appear to condemn nearly every discriminatory denial of an opportunity as a violation of autonomy and thus to legitimize nearly every application of antidiscrimination law according to Raz’s harm principle. But before endorsing this account of antidiscrimination law, we must ensure that it does not unduly burden employers and purveyors of public accommodations with duties detrimental to their own autonomy, duties whose fulfillment precludes engagement in activities and relationships essential for the autonomous life. If a particular application of antidiscrimination law deprives the perpetrator of more autonomy than it secures for the victim, if, say, the victim is relatively affluent and the perpetrator is engaged in a valuable activity or relationship that is particularly vulnerable to governmental interference, then Raz’s harm principle will condemn the application as illegitimate.

It does indeed seem that many employers and purveyors of public accommodations engage in precisely these kinds of delicate activities and relationships and, consequently, that they will often have to compromise their own autonomy to honor the autonomy-based duties imposed by antidiscrimination law, a potential sacrifice that casts serious doubt on the legitimacy of many of the legislation’s applications. The sensitivity of an activity or a relationship to governmental interference depends largely on the particular meaning with which broader social practice imbues it. For, at least in contemporary societies, the activities and relationships that constitute our “comprehensive goals,” those goals around which we organize our lives, must be based on what Raz calls “social forms,” “forms of behaviour which are in fact widely practised in . . . society.”¹⁰⁷ Raz identifies two reasons for the dependence of comprehensive goals on social forms.

¹⁰⁷ Ibid., 308.

First, “some comprehensive goals require social institutions for their very possibility. One cannot pursue a legal career except in a society governed by law, one cannot practice medicine except in a society in which such a practice is recognized.”¹⁰⁸ A second, and subtler, reason why comprehensive goals depend on social forms is that one can learn how to participate in many activities and relationships not through “explicit deliberation,” but only through “habituation.”¹⁰⁹ Some activities and relationships, in other words, are simply “too dense to allow explicit description or learning, they can be learnt only by experience.”¹¹⁰ One cannot, for example, learn how to become a good friend just by reading a manual, listening to a lecture, or even observing others’ friendships in action. Rather, one must participate in friendships for oneself, gradually acquiring an intuitive sense of precisely what kinds of reactions and commitments the relationships demand. It would be impossible to acquire this sense within a single lifetime if society did not already recognize and support friendships and the kinds of relations that constitute them—if, that is, there were no social form of friendship.¹¹¹

To say that activities and relationships depend on social forms is to say that they are “socially defined and determined pursuits and activities.”¹¹² The activities and relationships in which we engage, in other words, derive their meaning and significance from their social environment. The kind of social environment that characterizes contemporary western liberal democracies has endowed some of our activities and relationships with a meaning that renders them acutely vulnerable to governmental interference. More specifically, by treating autonomy as a constituent element of the good life, that environment conditions the value of certain activities and relationships on their being pursued freely—not according to some external set of guidelines, but rather

¹⁰⁸ Ibid., 310.

¹⁰⁹ Ibid., 311.

¹¹⁰ Ibid., 312.

¹¹¹ Ibid.

¹¹² Ibid., 309.

according to their participants' own wishes and inclinations. As John Gardner explains, some activities and relationships "have come to be identified . . . with spontaneity and self-expression," a development that promotes autonomy, since "[a] society in which every path one can choose is hedged about with rules and regulations is not a society which is particularly conducive to personal autonomy."¹¹³ Autonomous persons must enjoy at least some opportunities for relatively unfettered expression, and contemporary society furnishes them by fostering activities and relationships that thrive on such creativity.

Because social practice has designated these activities and relationships as necessarily spontaneous and expressive, they cannot readily withstand external interference. They are, to borrow Gardner's term, "direction-sensitive"—especially likely "to be distorted or damaged by the fact that they have been directed from outside, engineered by some non-participant with influence over the participants."¹¹⁴ Among all the external forces that might imperil direction-sensitive activities and relationships, the government in general and the coercive power of the law in particular pose particularly serious threats. For not only will the law tend to "taint" particular instances of the activities and relationships by regulating them so intrusively that their participants come to view themselves as manipulated, passive bystanders rather than spontaneous, active creators; it will also tend to imbue the social forms on which the activities and relationships depend with a symbolic meaning such that they "come to be associated with directive interventions" rather than with the "spontaneity and self-expression which they previously stood for," evolving into new, more regimented activities and relationships with considerably less value for autonomous persons.¹¹⁵ By thus compromising direction-

¹¹³ Gardner, "Private Activities," 153, 154.

¹¹⁴ *Ibid.*, 153.

¹¹⁵ *Ibid.*

sensitive activities and relationships, legal coercion can impose a considerable cost on its targets' autonomy.

Antidiscrimination law seems to epitomize the kind of intrusive legal regulation that entails such a cost. For by compelling employers and purveyors of public accommodations to discharge their autonomy-based duties, to provide opportunities to the racial minorities and women with whom they interact, antidiscrimination law often encroaches on—and sometimes even upends—some of the direction-sensitive activities and relationships in which they are engaged and that greatly enhance their autonomy. More specifically, the legislation requires activities and relationships in the workplace and other public venues to conform to certain governmentally prescribed criteria, supplanting the spontaneous, expressive choices of the persons participating in those activities and relationships with the government's homogenizing ones. However enlightened the government's choices might be, they preclude the authenticity on which direction-sensitive activities and relationships depend and thus extinguish valuable options. We must therefore ensure that applications of antidiscrimination law do not spawn "situations in which the enforcement of . . . autonomy-based duties will be counterproductive from the perspective of personal autonomy itself"—that the legislation does not, in other words, cause "serious institutional harm, depleting or skewing the society's general stock of autonomy-enhancing social forms, leaving too little space for truly spontaneous relationships, destroying more personal autonomy than it creates."¹¹⁶ Only if antidiscrimination law offsets this harm to direction-sensitive activities and relationships by significantly augmenting the autonomy of racial minorities and women will Raz's version of the harm principle legitimize the legislation.

¹¹⁶ Ibid., 155.

A proponent of Raz’s autonomy-based political morality might object that the preceding line of criticism romanticizes the workplace and the other public venues governed by antidiscrimination law, presupposing images of the employer-employee and merchant-customer relationships that belie actual social practice. Indeed, the activities and relationships that obtain in the workplace and similar settings seem to be not direction-sensitive, but rather nearly impervious to governmental interference. According to Gardner, only intensely intimate relationships—like those between spouses, partners, and friends—and profoundly personal activities—like voting for public officials or writing a novel—are direction-sensitive.¹¹⁷ Nor, Gardner argues, should we automatically deem an activity or relationship direction-sensitive simply because it inhabits a realm of social life traditionally regarded as “private,” for even some putatively “private” relationships—like those among the members of some social clubs—appear neither to subsist on spontaneity and self-expression nor to respond so aversely to governmental interference. If even these latter relationships are not direction-sensitive, then it is difficult to see why the employer-employee and merchant-customer relationships would be damaged when subjected to the requirements of antidiscrimination law.

Although this objection worries that we might pronounce too many activities and relationships direction-sensitive, exaggerating their vulnerability to external forces like governmental interference, it actually risks committing the opposite error. For, as we discovered through the analysis of freedom of association in Chapter 2, persons can associate with one another for multiple reasons, a complex of motives that can lend a multi-faceted character to their activities and relationships. The members of an economic firm, for example, might seek through their association not only to further their more prosaic, pecuniary interests, but also to pursue various valuable options like friendship

¹¹⁷ Ibid.

and the community of undertaking a shared enterprise. To be sure, not all persons found or join businesses to realize such lofty goals, but the many who do will seamlessly integrate into their overtly economic transactions activities and relationships that are quintessentially direction-sensitive. Lest it blindly impinge on its targets' autonomy, antidiscrimination law must appreciate these nuances in people's social interactions, identifying those valuable workplace activities and relationships it will tend to undermine and staying its hand when it cannot offset such harm with comparable benefits to the autonomy of racial minorities and women.

Even if some of the activities and relationships in which employers and purveyors of public accommodations participate are indeed direction-sensitive, we still might resist insulating them from the requirements of antidiscrimination law. After all, if a person can sustain an activity or relationship—however direction-sensitive—only by discriminating against racial minorities or women and hence shirking his autonomy-based duties, then in what sense does the activity or relationship remain a *valuable* one? We might suppose, in other words, that, by employing vicious means like discrimination to realize an otherwise valuable option, a person thereby adulterates the option indelibly, annulling any contribution it would have otherwise made to his autonomy. On this view, antidiscrimination law need not balk at regulating a direction-sensitive activity or relationship underwritten by discrimination, but rather may dissolve it to compel its participants to honor their autonomy-based duties.

Were we to endorse this conclusion, however, we would license the government virtually to extirpate direction-sensitive activities and relationships. For nearly every such activity or relationship blends good and evil, virtue and vice, belying the putatively rigid dichotomy between valuable and corrupt options. Indeed, although some corrupt direction-sensitive activities and relationships might lack even a modicum of value, most

instantiate important goods and consequently enhance their participants' autonomy. The government thus cannot reflexively regulate every corrupt direction-sensitive activity or relationship without thereby insinuating itself ubiquitously into people's lives and eviscerating direction-sensitive social forms. The government, in other words, must either tolerate at least some "admittedly imperfect spontaneous and self-expressive relationships [and activities]" or risk eliminating all such relationships and activities—valuable and corrupt alike.¹¹⁸ Because direction-sensitive activities and relationships constitute such a significant portion of the valuable options available in contemporary western liberal democracies, a government that values autonomy will pursue the former course rather than the latter one. Similarly, antidiscrimination law must abide considerable corruption of direction-sensitive activities and relationships if it is to respect autonomy and heed Raz's harm principle.

But even if we suppose that discriminatory direction-sensitive activities and relationships actually have no value for their participants, that they forfeit all their worth once they violate the autonomy of racial minorities and women, it still does not follow that the government may subvert them to enforce their participants' autonomy-based duties. For, according to Raz's autonomy-based morality, the government should concern itself not just with the upshots of the choices its citizens make, but also with the manner in which its citizens make them—not just with the options its citizens pursue, but also with the considerations that lead its citizens to pursue them. More specifically, the government should typically defer to its citizens on "matters (say, the choice of one's friends) where it is at least as important that people should choose for themselves as that they should choose wisely."¹¹⁹ Were the government instead to deny its citizens this latitude to err, regularly preempting some of their most intimate decisions, it would not

¹¹⁸ Ibid., 154.

¹¹⁹ Raz, "Liberty and Trust," 119-20.

simply eradicate corrupt options, but also enervate the capacities its citizens require to exercise any autonomy at all. Because we tend to manifest these capacities most authentically through our direction-sensitive activities and relationships, the government must be especially cautious when attempting to ameliorate even the most defective instantiations of such activities and relationships, refraining from imposing autonomy-based duties that would severely constrain the participants' autonomy. Though admittedly unpalatable, this forbearance of undeniably immoral activities and relationships is an inevitable implication of Raz's commitment to autonomy, a commitment that, while perfectionist, nevertheless insulates even some of our corrupt choices from governmental interference.

Of course, to insist that the government afford corrupt direction-sensitive activities and relationships significant leeway is not to forbid it to regulate them altogether. On the contrary, when such activities and relationships diminish the autonomy of third parties more than they benefit the autonomy of their participants, Raz's harm principle will permit the government to disrupt them. Many discriminatory direction-sensitive activities and relationships—especially those involving employers and purveyors of public accommodations—will indeed have such deleterious effects, diminishing the autonomy of racial minorities and women who seek to access the various opportunities provided by their participants. But because Raz conceives of his version of the harm principle merely as a means of realizing his broader autonomy-promoting goals, the government may not automatically assume that any benefits a particular application of antidiscrimination law might secure for racial minorities and women exceed the costs it might exact from the participants in discriminatory direction-sensitive activities and relationships. Rather, the government must weigh the former effects against the latter ones in order to determine whether the application will, on balance, promote more

autonomy than it curbs. As we have seen, this calculus will actually militate against many applications of antidiscrimination law, particularly where racial minorities and women enjoy choice sets sufficiently robust to withstand occasional denials of opportunities.

The Antiperfectionist Implications of Raz's Perfectionist Commitment to Autonomy

Despite his professed perfectionism, despite even his authorizing the government to coerce its citizens to discharge their autonomy-based duties, Raz struggles to justify certain core applications of antidiscrimination law. This difficulty owes largely to his singular focus on the value of autonomy, which overshadows other, equally integral elements of human flourishing. To be sure, Raz's conception of the good life undoubtedly recognizes values beyond autonomy; hence his value pluralism. And yet, these other values figure not at all in either his account of the harm principle in particular or his defense of liberalism more generally, an omission that bespeaks an aversion to coercive moral paternalism nearly as intense as any antiperfectionist's. Indeed, it is his exaltation of autonomy over all other values, his portrayal of the good life as, first and foremost, the autonomous one, that prompts Raz to balk at full-blown perfectionism, categorically enjoining the government from employing coercion to inculcate among its citizens character traits other than those immediately associated with autonomy. As we now have learned, however, a comprehensive defense of antidiscrimination law requires not this limited embrace of perfectionism, but rather a greater willingness to coercively cultivate moral character—in particular, a willingness to compel employers and purveyors of public accommodations to treat racial minorities and women in a certain manner even when they must abrogate some of their own activities and relationships and thus sacrifice some of their own autonomy to do so. Raz's autonomy-based political morality cannot underwrite this more moralistic mission.

Insofar as it privileges personal autonomy over other values, Raz's political morality resembles James Griffin's theory of human rights—and suffers from similar vulnerabilities.¹²⁰ Griffin attempts to develop an account of human rights “centered on the notion of agency,” an account that conceives of human rights as “protections of our agency—what one might call our personhood.”¹²¹ According to Griffin, to function as full agents and to realize our full personhood, we must enjoy both autonomy—the ability to make our own decisions for ourselves—and liberty—the freedom to act on those decisions, once made, without interference by others.¹²² All sound human rights protections, Griffin concludes, seek to secure these prerequisites of agency, a function that need invoke no values other than those of agency, autonomy, and liberty.¹²³

As John Tasioulas rightly notes, however, Griffin's rather reductive account of human rights overlooks “universal human interests beyond autonomy and liberty” that human rights might safeguard, thereby circumscribing considerably the range of available justifications for human-rights claims.¹²⁴ More specifically, by espousing a “dualist theory of the values that justify human rights” rather than a “pluralist” one, Griffin not only neglects the inherent interdependence between autonomy and liberty, on the one hand, and all the other elements of the good life, on the other, but also forecloses alternative justifications for human-rights claims that do not advert to the values of autonomy and liberty.¹²⁵ Griffin thus ends up propounding an emaciated account of human rights that captures only one facet of their complex relationship with human flourishing.

¹²⁰ In almost all other respects, however, Raz's account of autonomy bears only a superficial resemblance to Griffin's.

¹²¹ James Griffin, “Discrepancies Between the Best Philosophical Account of Human Rights and the International Law of Human Rights,” *Proceedings of the Aristotelian Society* 101 (2001): 4. See also James Griffin, “First Steps in an Account of Human Rights,” *European Journal of Philosophy* 9 (2001): 306-27.

¹²² Griffin, “Discrepancies,” 4-5.

¹²³ *Ibid.*, 7.

¹²⁴ John Tasioulas, “Human Rights, Universality and the Values of Personhood: Retracing Griffin's Steps,” *European Journal of Philosophy* 10 (2002): 91.

¹²⁵ *Ibid.*, 93-94.

Raz likewise disregards values other than autonomy that should feature prominently in any sound political morality. This neglect, in turn, hampers his defense of antidiscrimination law, suppressing principles that plausibly might ground many of the legislation's applications. Only by invoking these other principles, only by sometimes subordinating autonomy to other perfectionist goals, can Raz justify abridging all direction-sensitive activities and relationships to forestall discrimination. For unless the government may at least occasionally violate its citizens' personal autonomy to cultivate their moral characters, it will have to tolerate many discriminatory instances of such activities and relationships—namely, those instances that benefit their participants' autonomy more than they diminish that of third parties. However perfectionist, a commitment to autonomy alone cannot legitimate all applications of antidiscrimination law.

IV. CONCLUSION

In expounding his political morality, Raz repeatedly repudiates antiperfectionism, unabashedly adverting to a conception of the good life and explicitly enjoining the government to promote it. And yet, when we attempt to apply that political morality to the particular issue of discrimination based on race or sex, we discover that it actually shares many of the features—and some of the pitfalls—of its antiperfectionist counterparts. Most notably, it turns out that Raz espouses a version of the harm principle that, while undoubtedly rooted in a conception of the good life, nevertheless circumscribes the government's regulatory agenda in general and nondiscrimination program in particular nearly as severely as would an antiperfectionist one. For even when we construe harm in terms of the value of autonomy, understood as an integral component of human flourishing, many instances of discrimination will prove harmless, and the harm principle will thus condemn many applications of antidiscrimination law as

illegitimate—especially as racism and sexism wane and the material circumstances of racial minorities and women improve.

Nor can Raz avert this failure by incorporating into his theory a more moralistic conception of harm, one according to which shirking autonomy-based duties is itself harmful—regardless of the material circumstances of the intended beneficiaries of those duties. For, despite its perfectionist provenance, Raz’s harm principle seeks to safeguard not just the autonomy of the victims of discrimination, but also that of the perpetrators, and thus will condemn those applications of antidiscrimination law that curtail the latter more than they augment the former. In particular, because many employers and purveyors of public accommodations participate in direction-sensitive activities and relationships, activities and relationships acutely vulnerable to governmental interference, the government will often be unable to compel them to discharge their autonomy-based duties without thereby rupturing those activities and relationships and abridging their autonomy. Unless offset by commensurate benefits to the autonomy of the victims, these costs to the autonomy of the perpetrators will render antidiscrimination law illegitimate according to Raz’s harm principle.

Raz’s perfectionist political morality thus fails to legitimize many core applications of antidiscrimination law, a failure that stems from its nearly exclusive focus on autonomy and its elevation of that value as the government’s paramount concern. Only by invoking values other than autonomy can we condemn all instances of discrimination, and only by occasionally infringing people’s autonomy to cultivate their moral characters can we vindicate all applications of antidiscrimination law.

Part II

**A Perfectionist Account of
Antidiscrimination Law**

Chapter 5

Discrimination and Respect

I. BEYOND LIBERAL ACCOUNTS OF ANTIDISCRIMINATION LAW

Three of the most prominent contemporary liberal doctrines of political legitimacy fail to legitimize certain core applications of antidiscrimination law, applications that proscribe forms of behavior we deem especially objectionable and that epitomize the kind of governmental regulation we deem obviously legitimate. Though perhaps somewhat surprising, this failure was foreshadowed by the three doctrines' common approach to discrimination based on race or sex. For each of the doctrines condemned discrimination not as an intrinsic wrong that blemishes the moral characters of its perpetrators, but merely as a contingent one that tends to have various deleterious consequences for its victims. To be sure, any plausible account of discrimination will acknowledge and duly respond to these consequences, authorizing the government to implement policies to mitigate them. But unless such an account also looks beyond the contingent consequences of discrimination, attending to some of the conduct's inherent characteristics, it will succumb to precisely the same shortcoming that afflicted each of the three liberal doctrines of political legitimacy considered in the previous three chapters—namely, an ability to underwrite, at most, only a sporadic, discontinuous antidiscrimination regime, one that extends its protections only to particularly disadvantaged members of historically oppressed social groups and only in particularly prejudiced social milieus.¹ An account of

¹ Of course, we might think that antidiscrimination regimes *should* be sporadic and discontinuous in other respects. For example, some might lament the formal symmetry of antidiscrimination law, the legislation's willingness to protect, say, men as readily as it would women from discrimination based on sex. Others might decry the expansion of antidiscrimination law to encompass discrimination based on traits like religion, which seem to bear little resemblance to the core ones of race and sex. But regardless of the validity of such critiques, we tend to laud antidiscrimination law's willingness to protect, say, affluent black people as readily as it would indigent ones from discrimination based on race—and to do so in tolerant social contexts as readily as it would amid pervasive prejudice. It is these latter, commonly accepted kinds

discrimination that focuses exclusively on the effects such behavior tends to have for its victims cannot but be confounded when those effects fail to materialize.

Nor is this focus on the consequences of discrimination some incidental feature of the three liberal doctrines of political legitimacy, a mere infelicity we can excise from the doctrines without altering their essential natures. On the contrary, such a focus is dictated by the distinctively liberal principles underlying each of the doctrines. Observing the strictures of antiperfectionism and public reason, both Rawls's political liberalism and an antiperfectionist version of the harm principle forbid the government to promote any particular conception of the good life and to cultivate its citizens' moral characters. And while Raz unequivocally repudiates antiperfectionism and its doctrine of public reason, he nevertheless espouses an understanding of the good life rather congenial to at least some of those values' practical implications, an understanding that privileges personal autonomy and hence severely circumscribes the government's authority to employ the law to inculcate virtue and curb vice. Whether perfectionist or antiperfectionist in provenance, an overriding commitment to personal autonomy and a corresponding aversion to coercive moral paternalism unite the three doctrines of political legitimacy and distinguish their contemporary brand of liberalism.

The doctrines would contravene these foundational liberal tenets were they to reconstrue antidiscrimination law as a measure for reforming the moral characters of perpetrators as well as for ameliorating the material privation of victims. For by endowing antidiscrimination law with the former purpose, the doctrines would invite the government not only to adjudge the nature of the good life, but also to restrict its citizens' personal autonomy in the name of moral virtue—to impose, in other words, precisely the kinds of restrictions that antiperfectionism, public reason, and even an overriding

of uniformity and continuity in antidiscrimination law that the three liberal doctrines of political legitimacy cannot justify.

perfectionist commitment to personal autonomy preclude. No doctrine of political legitimacy can so endorse legal moralism and still profess its adherence to contemporary liberalism.

It might be objected that this conclusion rests on a false dichotomy, whereby an account of discrimination must focus either on the consequences the conduct has for its victims, in which case the account will remain antiperfectionist but fail to justify certain core applications of antidiscrimination law, or on the moral characters of discriminators, in which case the account will justify all applications of antidiscrimination law but become perfectionist. In fact, many liberals will favor a third position, according to which discrimination is wrong in virtue of neither the material consequences it has for its victims nor the characters of its perpetrators, but rather the fact that it violates its victims' rights. It would, of course, be question-begging simply to assert the existence of a right not to be discriminated against; we must ground such a right in still more fundamental moral principles, ones that explain why discrimination is the kind of wrong that justifies protecting its victims with a right and imposing on employers and purveyors of public accommodations a correlative duty to eschew the conduct.

One kind of principle might ground a right against discrimination in the victims' interests or well-being, which discrimination might be thought to compromise in various ways. But we need only state this possibility to recognize the resemblance it bears to the accounts of discrimination derived from Feinberg's and Raz's versions of the harm principle in the previous two chapters. Just like those accounts, a welfarist rights-based account of discrimination would fail to justify many applications of antidiscrimination law, since it would similarly focus on the adverse consequences discrimination has for its victims and condemn the conduct only insofar as those consequences actually

materialized. We cannot rescue harm-based accounts of discrimination simply by recasting them in rights-based language.

Another kind of principle might ground a right against discrimination not in the victims' well-being, but in their personhood. On this view, the right against discrimination would be but one instantiation of a more general Kantian right to be treated with the respect to which all persons are entitled, and discrimination—the violation of that right—would constitute an intrinsic wrong, one wholly independent of any material consequences for the victims. This chapter will develop such an account of the wrongfulness of discrimination and thereby provide the basis for a right not to be discriminated against. But once we have developed a Kantian account of the wrongfulness of discrimination, we must ask whether it remains antiperfectionist, whether it specifies the wrongfulness of discrimination—and the right not to be discriminated against—independently of discriminators' moral characters. The next chapter will argue that a Kantian account of discrimination does indeed turn out to be perfectionist—in particular, that we cannot understand discrimination as a kind of disrespect without attending to the moral characters of discriminators and that we consequently cannot legitimize antidiscrimination law without espousing perfectionism.

Before proceeding with these arguments, it is important to concede that proponents of the doctrines of political legitimacy considered in the previous three chapters might plausibly continue to scorn a perfectionist approach, refusing to jettison the doctrines even to rescue certain core applications of antidiscrimination law. In particular, they might privilege their allegiance to the doctrines over any partialities they might have for particular applications of antidiscrimination law—however integral such applications might be to the mission of contemporary antidiscrimination regimes. After all, the proponents presumably espoused their favored doctrines in the first place only

after reflecting deeply on the proper relationship between the state and its citizens, reflections, we might think, that no independent intuition about the immorality of a particular form of individual conduct or the legitimacy of a particular form of governmental regulation should trump.

The intelligibility of this response highlights the largely ad hoc nature of this thesis's critique of the three contemporary liberal doctrines of political legitimacy. Rather than launch a frontal assault on the fundamental liberal tenets underlying the doctrines, that critique instead sought to impugn contemporary liberalism indirectly by exposing its incompatibility with central aspects of an especially cherished political institution, one that we commonly assume embodies some of our most confident moral judgments. Admittedly, then, contemporary liberalism remains a perfectly coherent ideology even in the aftermath of this oblique offensive. But its luster has now faded considerably. For we can no longer trust contemporary liberalism to accommodate our wholesale condemnation of discrimination or our unconditional support for antidiscrimination law; either some of our ideological predilections or some of our firmly held first-order moral intuitions must yield. Assuming that our fidelity to contemporary antidiscrimination regimes is not misplaced, assuming that the government is right to proscribe even those instances of discrimination that do not significantly harm their victims, we should seek an alternative, comprehensively perfectionist account of discrimination, one that focuses on the intrinsic wrongfulness of the conduct and the moral characters of its perpetrators and thus transcends many of the vicissitudes of social and economic practice.

This chapter will begin to develop such an account. After reconciling our intuition that discrimination is always wrong, irrespective of its consequences, with the conduct's historically contingent nature, this chapter will argue that the wrongfulness of discrimination inheres in the reasoning of discriminators. More specifically,

discrimination is always wrong because acting on discriminatory intentions is always wrong. It is important to clarify at the outset that this chapter uses the term “intention” neither in the sense in which it is often used in tort law, according to which one acts “intentionally” so long as one is merely aware of one’s actions and their consequences, nor in a sense that would equate intentions with “motives,” the ultimate reasons for which people act. Rather, for purposes of this chapter, one acts on a discriminatory intention when one takes another person’s race or sex as a reason to treat her less favorably than one would treat people of other races or the other sex—whether such treatment is a means to some further end or an end in itself. When one takes another person’s race or sex as a reason to treat her less favorably, it will be argued, one acts immorally because one fails to treat her with the respect to which she is entitled in virtue of her personhood, to properly value her humanity in one’s deliberations. This chapter will distinguish this account of discrimination as a form of disrespect from other accounts that locate the conduct’s wrongfulness in the disrespectful social meanings conveyed by discriminatory acts rather than the disrespectful intentions of discriminators.

It will be left to the next chapter to show that, *pace* traditional renditions of Kantian moral theory, the moral wrongfulness of discrimination and the moral characters of discriminators are inextricably linked, that we cannot adopt an intentionalist, respect-based account of discrimination unless we are willing to assess moral character—even if it is true, as many Kantian theorists believe, that we can generally assess people’s intentions without considering their characters. The next chapter will also show that we cannot understand antidiscrimination law as a measure for enforcing a Kantian right to be treated with respect unless we are willing to espouse full-blown perfectionism.

II. DISCRIMINATION AND HISTORICAL CONTINGENCY

In criticizing contemporary liberal doctrines of political legitimacy, Part I of this thesis proceeded on the assumption that the government should never tolerate discrimination based on race or sex. But is this assumption really warranted? Does every instance of the conduct truly merit prohibition, or may we content ourselves after all with the sporadic, discontinuous antidiscrimination regimes licensed by the three liberal doctrines of political legitimacy? While we might initially balk at the latter conclusion, we should consider the possibility that our contrary intuition, our belief that the government must proscribe all discriminatory acts, derives not from our philosophical reflections about the moral wrongfulness of discrimination, but from our empirical observations of actual, concrete manifestations of more generic race- and sex-based injustice. Alan Wertheimer, for example, suggests that our present intuitions about discrimination

may well be closely tied to the actual social consequences of discrimination, or to the history of those consequences. If whites and blacks had relatively equal wealth, income, longevity, education, and so forth, it is entirely possible that we would not in fact regard isolated acts of racial discrimination as seriously wrong.²

According to this view (which Wertheimer does not definitively endorse), discrimination is sufficiently wrong to warrant legal prohibition only insofar as it engenders or compounds the disadvantage of racial minorities and women, “contribut[ing] to a pattern of social and economic subordination.”³ It is, in other words, the systemic upshot of discrimination, not its intrinsic immorality, that necessitates antidiscrimination law.

In decrying discrimination for its contingent, systemic consequences rather than for its other, intrinsic features, scholars appear to seek a historically sensitive account of the conduct, one that roots the wrongfulness of discrimination in the special histories of race and sex in contemporary western societies. Absent these peculiar histories, the

² Alan Wertheimer, “Reflections on Discrimination,” *San Diego Law Review* 43 (2006): 952 (emphasis omitted).

³ *Ibid.*

argument goes, we would tend to ignore discrimination based on race or sex, largely dismissing it as a mere aberration that merited little moral censure—let alone a legal remedy. We would, in other words, regard such discrimination much as many people now regard discrimination based on other arbitrary characteristics like eye color.⁴ Indeed, according to Richard Wasserstrom, it is only because our societies, over time, have fixated on race and sex as “socially important categories” that we so emphatically denounce discrimination based on those properties and so wholeheartedly endorse comprehensive antidiscrimination law.⁵ And it seems that this historical legacy elicits such vehement moral condemnation only because it has culminated in systematic disadvantage for racial minorities and women, yielding a set of profoundly adverse consequences that alerts our moral conscious to forms of conduct we would otherwise ignore.

But must we really focus on the consequences of discrimination to take the special histories of race and sex in our societies seriously? By insisting that the wrongfulness of discrimination inheres not only in the consequences it has for its victims, but also in its other, intrinsic features, do we necessarily embrace a dubious ahistorical moral universalism according to which “moral truths such as the wrongfulness of certain kinds of discrimination have always been with us, and . . . it merely took us until the second half of the twentieth century to discover them”?⁶ *Pace* theorists like Wasserstrom, it seems we can assess features of discriminatory acts other than their effects without thereby having to ignore historical context altogether. For even if history does not determine the morality of discrimination, it still might affect the conduct’s suitability for legal proscription. We might, for example, deem both discrimination based on race or sex and discrimination based on eye color equally wrongful but nevertheless tolerate the latter

⁴ Richard Wasserstrom, “Racism, Sexism, and Preferential Treatment,” *UCLA Law Review* 24 (1976): 586.

⁵ *Ibid.*, 584, 586, 589-90.

⁶ John Gardner, “On the Ground of Her Sex(uality),” *Oxford Journal of Legal Studies* 18 (1998): 177.

conduct because, as a matter of historical contingency, it tends not to be nearly so prevalent or insidious as the former.

But even if we concede that history does determine the morality of various forms of discrimination, it does not follow ineluctably that the morality of discriminatory acts depends on their effects. After all, the peculiar history of racism and sexism might shape not only the consequentialist upshot of discrimination based on race or sex, but also its other features. Perhaps that history dictates the social meaning of race- and sex-based discriminatory acts, necessarily rendering such acts “demeaning” toward their victims;⁷ or perhaps it indelibly adulterates the moral reasoning of those persons who discriminate based on race or sex, transmuting morally indifferent, or generically immoral, intentions into disrespectful ones.⁸ While linking morality inextricably with history, neither of these accounts of discrimination indicts the conduct merely for the consequences it has for its victims. On the contrary, each account appears to denounce every instance of discrimination—even those instances that leave their victims more or less unscathed. We thus need not focus on consequences to situate discrimination in its historical context.

To be sure, both of the preceding rejoinders to the historicist challenge presuppose the truth of an alternative account of discrimination like the two cited in the preceding paragraph, an account that locates the wrongfulness of discrimination in features other than its effects. Some theorists, however, simply do not share the seemingly widespread intuition that discrimination is wrong (or at least sufficiently wrong to merit legal prohibition) even when it causes no significant harm.⁹ But if we can successfully defend one of the two alternative accounts of discrimination, we will have condemned the conduct even when it creates no such disadvantage and thus will have vindicated the

⁷ Such a view is propounded by Deborah Hellman in *When Is Discrimination Wrong?* (Cambridge, Mass.: Harvard University Press, 2008) and will be examined more closely below.

⁸ This is the account of discrimination based on race or sex that this chapter will ultimately defend.

⁹ See, e.g., Alan H. Goldman, “Affirmative Action,” *Philosophy & Public Affairs* 5 (1976): 192-93.

intuition that every discriminatory act—whether against the indigent or the affluent—is wrong. We will have done so, moreover, without neglecting the special histories of race and sex in our societies.

III. BEYOND THE “VICTIM PERSPECTIVE” AND SYSTEMIC DISCRIMINATION

Contemporary antidiscrimination law itself seems to suggest an alternative focus we might adopt when evaluating discrimination. For rather than considering only the consequences discrimination has for its victims, the legislation attends primarily to the reasoning of the conduct’s perpetrators. U.K. antidiscrimination law, for example, originally defined direct discrimination in terms of the “grounds” on which the perpetrator decides to treat a person “less favourably” than others.¹⁰ According to this definition, discrimination is a product not of the circumstances of its victims, but of the reasoning of its perpetrators.

To be sure, while antidiscrimination law *defines* discrimination in terms of the reasoning of perpetrators, it *prohibits* the conduct only in contexts where victims are likely to be deprived of important benefits and opportunities. More specifically, only after the legislation applies its definition of discrimination to the contexts of employment, education, housing, public accommodations, and so on does it create any legal duties not to discriminate.¹¹ Because the actual prohibitions against discrimination contemplate the effects the conduct has on its victims as well as the reasons for which its perpetrators act, we might be tempted to revisit an account of antidiscrimination law like Raz’s, a liberal-perfectionist account that permits the government to define the wrongfulness of discrimination in perfectionist terms yet forbids it to employ coercion to remedy the conduct unless the victims are also harmed. Perhaps, we might think, Raz was right after

¹⁰ See, e.g., Race Relations Act 1976, § 1(1)(a).

¹¹ See, e.g., *ibid.*, parts II and III.

all to conceive of antidiscrimination law as a fundamentally perfectionist measure that nonetheless attempts to comply with the liberal harm principle.

Before accepting this conclusion, however, we should consider an alternative explanation for why the legislation might proscribe discrimination only in contexts involving the provision of particularly important opportunities and benefits. Perhaps antidiscrimination law so circumscribes its purview not to observe the strictures of the harm principle, but to realize some pragmatic or institutional objective, such as curbing lawsuits. Were the legislation to forbid discrimination even in circumstances where no significant opportunities or benefits were at stake, the argument goes, it would inundate the courts with relatively trivial claims and thus distract their attention from more urgent grievances. Unlike Raz's account of antidiscrimination law, this one is purely perfectionist, tolerating certain harmless instances of discrimination in order to heed a prudential imperative rather than the strictures of the liberal harm principle. And on reflection, it does indeed seem that it is prudence rather than principle that prompts antidiscrimination law to restrict its prohibitions to contexts like employment and the provision of public accommodations. For to satisfy the harm principle, antidiscrimination law would have to not only focus exclusively on discriminatory deprivations of opportunities and benefits, but also distinguish between mere deprivations and truly harmful ones—a distinction that, as we saw in the previous two chapters, antidiscrimination law makes no effort to draw. At least at the foundational level of moral principle, then, antidiscrimination law appears to prioritize its reason-sensitive definition of discrimination over its particular context-sensitive prohibitions of the conduct, establishing the reasoning of discriminators rather than the circumstances of victims as its focal concern.

But by focusing on the perpetrators of discrimination rather than the victims, some theorists worry, antidiscrimination law might actually abet discrimination, insulating the conduct from the most effective governmental remedies. Alan Freeman, for example, laments antidiscrimination law's adoption of what he dubs the "perpetrator perspective," which "sees racial discrimination . . . as actions, or series of actions, inflicted on the victim by the perpetrator. The focus is more on what particular perpetrators have done or are doing to some victim than it is on the overall life situation of the victim class."¹² We can improve "the overall life situation" of discrimination's victims, Freeman argues, only if we define discrimination from the "victim perspective" and attend to "the conditions of actual social existence as a member of a perpetual underclass," conditions comprising "both the objective conditions of life—lack of jobs, lack of money, lack of housing—and the consciousness associated with those objective conditions—lack of choice and lack of human individuality in being forever perceived as a member of a group rather than as an individual."¹³ According to Freeman, to emulate antidiscrimination law in defining discrimination from the "perpetrator perspective" is to countenance the continued privation of racial minorities, to pursue the moral censure of discriminators at the expense of the individual well-being of victims.

Thus articulated, Freeman's critique of the "perpetrator perspective" seems to reflect certain distinctively institutional concerns. More specifically, Freeman seems to lament the institutional implications of enshrining the "perpetrator perspective" in law, of defining the legal offense of discrimination primarily in terms of the discriminator's reasoning. Perhaps willing to acknowledge that the immorality of discrimination stems at least in part from features of the discriminator's reasoning, he nevertheless seeks to downplay—or even obfuscate—that fact in law, lest "courts becom[e] bogged down in

¹² Alan David Freeman, "Legitimizing Racial Discrimination Through Antidiscrimination Law: A Critical Review of Supreme Court Doctrine," *Minnesota Law Review* 62 (1978): 1053.

¹³ *Ibid.*, 1052-53.

technicality or fail[] to recognize the very great moral importance of discrimination's effects upon those on the receiving end of it."¹⁴ But Freeman's critique also suggests a more radical, conceptual argument, according to which adopting the "perpetrator perspective" is not just institutionally imprudent, but morally impertinent, since, "morally speaking, only the *effects* of discriminatory decisions count in making them discriminatory."¹⁵ By focusing on the discriminator's reasoning, this view suggests, we neglect what should be the true object of our moral concern: the consequences discrimination has for its victims.

Construed in this second way, as a conceptual claim about the nature of discriminatory acts, the idea of the "victim perspective" seems to support a systemic, rather than an individualistic, account of discrimination. For by heeding the imperative of the "victim perspective" to view the victims of discrimination as members of certain disadvantaged social groups, we largely abstract from particular instances of discrimination, subsuming them under a broader social practice of subordination and disadvantage. A systemic account of discrimination, in other words, rebukes discriminators not for merely "taking a person's race or sex into account," but rather for

taking race and sex into account in a certain way, in the context of a specific set of institutional arrangements and a specific ideology which together create and maintain a *system* of unjust institutions and unwarranted beliefs and attitudes. That system is and has been one in which political, economic, and social power and advantage are concentrated in the hands of those who are white and male.¹⁶

Indeed, systemic accounts of discrimination like this one seem to direct their ire at the unjust "system" itself, the "correlation between being a member of a certain group and

¹⁴ Gardner, "On the Ground," 186.

¹⁵ *Ibid.*, 187.

¹⁶ Wasserstrom, "Racism, Sexism," 591-92. For a similar systemic account of discrimination, see T.M. Scanlon, *Moral Dimensions* (Cambridge, Mass.: Harvard University Press, 2008), 73. In the legal literature, see Owen Fiss, "Groups and the Equal Protection Clause," *Philosophy & Public Affairs* 5 (1976): 107-77; and Cass R. Sunstein, "The Anticaste Principle," *Michigan Law Review* 92 (1994): 2410-55

suffering a certain kind of disadvantage,” rather than at the discrete decisions of discriminators.¹⁷

In their relative disregard for questions of first-order, individual morality, systemic accounts of discrimination strongly resemble contemporary theories of social justice, which, as we saw in Chapter 2, tend to concern themselves primarily with the overall distribution of benefits and burdens in society and only secondarily (if at all) with the intrinsic morality of the actions that foster that distribution. In other words, like the approach theories of social justice adopt toward individual actions in general, the approach systemic accounts of discrimination adopt toward individual discriminatory acts in particular is “top-down” rather than “bottom-up”: rather than directly promulgating norms to govern the treatment of racial minorities and women by employers and other relevantly situated persons, systemic accounts of discrimination first posit an overall distribution of benefits and burdens that society must realize and then indict discriminatory acts for upsetting that pattern, for disadvantaging racial minorities and women vis-à-vis their compatriots.¹⁸ According to systemic theories, then, discrimination is a purely derivative wrong, one whose wrongfulness owes solely to its connection to a broader, more generic social injustice.

Though coherent, systemic accounts of discrimination do seem to undermine antidiscrimination law’s *raison d’être*. For once discrimination no longer contributes to race- and sex-based disadvantage, systemic theories must dismiss antidiscrimination law as superfluous. Systemic theories, in other words, assign antidiscrimination law a merely provisional mandate, countenancing the legislation only insofar as discrimination continues to diminish the socioeconomic status of racial minorities and women.

¹⁷ Matt Cavanagh, *Against Equality of Opportunity* (Oxford: Oxford University Press, 2003), 200.

¹⁸ *Ibid.*, 198-99.

More fundamentally, by espousing a systemic account of discrimination, we belie the ordinary meaning of the term “discrimination.” Whereas systemic accounts label as “discrimination” all race- and sex-based disadvantage as such, irrespective of its origin, we more commonly use the term to denote only a particular form of intentional conduct—namely, one’s decision to treat a person differently because of her race or sex.¹⁹ To be sure, this criticism is largely a semantic one; it is perfectly natural for proponents of systemic accounts to bend linguistic meaning in order to harness the rhetorical force of the term “discrimination.” But in doing so, they appear to conflate two distinct moral concerns.

Systemic accounts of discrimination also belie the structure of contemporary antidiscrimination law. Rather than reflecting any general distributive purpose, “[a]nti-discrimination laws are commonly structured in such a way as to suggest that discriminators have committed a personal wrong against their victims, akin to a tort.”²⁰ In particular, just like tort law, antidiscrimination law generally requires individual victims to lodge claims directly against individual perpetrators and compels individual perpetrators directly to compensate individual victims. Systemic accounts cannot explain these features, but must instead dismiss them as inconvenient impediments to the broader mission of achieving distributive justice for racial minorities and women. Not systemic patterns, but individual actions—these are the cardinal concern of antidiscrimination law, a concern that directs our attention to the perpetrators of discrimination.

IV. ANTIDISCRIMINATION LAW, MOTIVES, AND INTENTIONS

Direct Discrimination

In developing an account of discrimination, we should take antidiscrimination law on its own terms, noting its explicit references to the reasoning of discriminators and

¹⁹ Ibid., 199.

²⁰ Sophia Moreau, “What Is Discrimination?,” *Philosophy & Public Affairs* 38 (2010): 145-46.

adopting its focus on discrete discriminatory acts. Recall that, according to the legislation, an employer or other relevantly situated person commits (direct) discrimination whenever he treats another person “less favourably” than he would treat others “on grounds of”²¹ (or “because of”²²) her race or sex. Antidiscrimination law thus defines discrimination not primarily in terms of the consequences it has for its victims, but rather in terms of the reasoning processes of its perpetrators.

But how, exactly, might race or sex manifest itself in a person’s reasoning? We can represent every instance of practical reasoning as a syllogism, a logical inference by an agent from various premises to a conclusion about what is to be done in a given set of circumstances. More specifically, a practical syllogism comprises both a major (“operative”) premise and at least one minor (“auxiliary”) premise.²³ Whereas an operative premise can motivate an agent, prompting him to adopt a “practical critical attitude” toward certain actions, auxiliary premises serve merely as “identifying reasons,” empirical facts that “help identify the act which there is reason to perform” given the agent’s belief in a particular operative reason.²⁴ Consider the following practical syllogism:

- (1) I should phone my mother on her birthday.
- (2) It is my mother’s birthday today.
- (3) I should phone my mother today.

The operative premise, premise (1), motivates the agent, identifying a goal he has reason to pursue. By contrast, the auxiliary premise, premise (2), is itself “motivationally

²¹ U.K. Race Relations Act 1976, § 1(1)(a).

²² Title VII of the U.S. Civil Rights Act of 1964. The U.K. Equality Act 2010, the most recent revision of the Race Relations Act 1976 and Sex Discrimination Act 1975, adopts the “because of” formulation of U.S. antidiscrimination law. In its Explanatory Memorandum, Parliament made clear that it considered the “because of” and “on grounds of” formulations to be substantively equivalent: “[T]his change in wording does not change the legal meaning of the definition, but rather is designed to make it more accessible to the ordinary user of the [Act].”

²³ Gardner, “On the Ground,” 180. For more on the distinction between operative and auxiliary reasons, see Joseph Raz, *Practical Reason and Norms*, 2nd ed. (Oxford: Oxford University Press, 1999), 33-35.

²⁴ Raz, *Practical Reason*, 33, 34.

inert,”²⁵ since it merely transfers premise (1)’s practical force to the syllogism’s conclusion (3) by identifying a fact relevant for realizing the goal prescribed by premise (1).²⁶ Whether explicit or implicit, this kind of syllogistic structure frames every intentional human action.

While race or sex can infect a practical syllogism through either the operative or auxiliary premise, legal prohibitions against direct discrimination must concern themselves with the former kind of taint if they are to track morality, as we suppose them to do. For only when race or sex informs an agent’s goal, bestowing motivational force on a particular action, must the agent ask himself whether he should still pursue the goal and perform the action—an inescapably moral question that is, by contrast, impertinent when directed at auxiliary premises, which are purely “informational” and thus “fall to be judged by epistemic standards alone.”²⁷ Imagine an employer who reasons as follows:

- (1) I must not hire black people.
- (2) This applicant is black.
- (3) I must not hire this applicant.

The employer contravenes a moral norm by acting on the conclusion (3) only if the operative premise, premise (1), is itself morally problematic; the auxiliary premise, premise (2), is not amenable to moral evaluation. To be sure, the operative premise might stem from myriad racist sources, including: brute “antipathy” toward black people or indifference to their welfare,²⁸ a self-serving ideology that rationalizes racial hierarchy,²⁹ or prejudiced beliefs about black people.³⁰ But irrespective of the operative premise’s precise provenance, the employer commits unlawful direct discrimination so long as he

²⁵ Gardner, “On the Ground,” 180.

²⁶ Raz, *Practical Reason*, 34-35.

²⁷ Gardner, “On the Ground,” 182.

²⁸ See, e.g., J.L.A. Garcia’s writings on racism, including “The Heart of Racism,” in *Race and Racism*, ed. Bernard Boxill (Oxford: Oxford University Press, 2001); “Current Conceptions of Racism: A Critical Examination of Some Recent Social Philosophy,” *Journal of Social Philosophy* 28 (1997): 5-42; and “Philosophical Analysis and the Moral Concept of Racism,” *Philosophy & Social Criticism* 25 (1999): 1-32.

²⁹ See, e.g., Tommie Shelby, “Is Racism in the ‘Heart’?,” *Journal of Social Philosophy* 33 (2002): 411-20.

³⁰ For one definition of “prejudice,” see Richard J. Arneson, “What Is Wrongful Discrimination?,” *San Diego Law Review* 43 (2006): 788.

permits it to influence his behavior by acting on the conclusion of the syllogism it initiates.

An employer commits unlawful direct discrimination not only if he is racist, discriminating against black people for its own sake, but also if he discriminates against black people as a means to some ulterior end. For instance, rather than having a visceral aversion to black people or explicitly regarding them as morally inferior, the employer might consider them unqualified to work in his firm, reasoning as follows:

- (1) I must not hire people who cannot make widgets.
- (2) Black people cannot make widgets.
- (3) I must not hire black people.
- (4) This applicant is black.
- (5) I must not hire this applicant.

Here, the employer uses race as a proxy for another trait, which, in turn, he deems an essential prerequisite for successful widget making.³¹ He thus discriminates against black people merely as a means to the further end of productive efficiency, evincing no necessarily racist attitudes or beliefs.³² And yet, we tend to condemn such conduct no less adamantly. For while race initially figures only in an auxiliary premise, premise (2), it is subsequently incorporated by the first operative premise, premise (1), into the conclusion of the first syllogism (3), which then becomes a new operative premise for the second syllogism (propositions (3)-(5)).³³ We thus censure the employer not only on epistemological grounds, but also on moral ones—not only for erroneously believing premise (2), but also for adopting premise (3) as an operative premise in his reasoning and acting on the discriminatory conclusion (5) it supports. Antidiscrimination law tracks this moral judgment by proscribing discrimination based on proxies and stereotypes as readily as it proscribes more overtly racist forms of discrimination.

³¹ For more on discrimination based on proxies and stereotypes, see Larry Alexander, “What Makes Wrongful Discrimination Wrong? Biases, Preferences, Stereotypes, and Proxies,” *University of Pennsylvania Law Review* 141 (1992): 167-71.

³² To be sure, the belief expressed in premise (2) might (but need not) ultimately owe to racism.

³³ Cf. Gardner, “On the Ground,” 182.

An employer or other relevantly situated person need not even espouse particular beliefs (racist or otherwise) about racial minorities or women in order to commit unlawful discrimination. Instead, he might simply acquiesce in the prejudiced predilections of his clientele. He might, in particular, include among his hiring criteria certain “reaction qualifications,” “abilities or characteristics which contribute to job effectiveness by causing or serving as the basis of the appropriate reaction in the recipients.”³⁴

Contemplating the potential reactions of his clientele to his hiring, say, a black employee, the employer might reason as follows:

- (1) I must not lose my current clientele.
- (2) My current clientele are racist and would cease patronizing my business were I to hire any black people.
- (3) I must not hire any black people.
- (4) This applicant is black.
- (5) I must not hire this applicant.

In this case, race infiltrates the employer’s reasoning neither through his own bigotry or racist ideology nor through a race-based proxy or stereotype, but rather through his clientele’s racism, sublimated by the profit-maximizing imperative of the market.

Nevertheless, just like his racist or stereotyping counterparts, the employer ultimately espouses a patently discriminatory operative premise (3) and acts on the equally discriminatory conclusion (5) it supports. At least according to the definition of discrimination enshrined in antidiscrimination law, the employer commits a discrete discriminatory act, one that does not simply reduce to his clientele’s prejudice. For it is ultimately the employer himself, not his clientele, who decides on this particular occasion to take another person’s race as a reason to treat her less favorably. To be sure, we might also condemn the employer for being complicit in his clientele’s prejudice. But this further offense consists not merely in making a race-based decision, but in doing so with

³⁴ Alan Wertheimer, “Jobs, Qualifications, and Preferences,” *Ethics* 94 (1983): 100. See also Wertheimer, “Reflections,” 946, 966.

a particular ulterior motivation—namely, with the aim of accommodating (for whatever reason) other people’s prejudice. Antidiscrimination law prescind from the ulterior motivations with which discriminators act, focusing instead exclusively on their more immediate reasons.

It is thus clear that antidiscrimination law defines direct discrimination in terms of the reasoning of discriminators. And while the term “reasoning” is itself ambiguous, by considering the structure of the practical syllogism, we have made significant progress toward identifying the precise mental feature of discriminators to which antidiscrimination law attends. We have seen, for example, that antidiscrimination law does not attend to the auxiliary premises of discriminators’ practical syllogisms, premises that are purely informational and thus, taken alone, do not bear moral scrutiny. Nor does antidiscrimination law attend *directly* to the operative premises of discriminators’ practical syllogisms. It does not, for instance, consider whether an employer forms the belief that he should not hire black people based on his own animus or prejudice, popular stereotypes, or his customers’ preferences. Rather, antidiscrimination law attends only *indirectly* to the operative premises—that is, only insofar as they authorize the logical inference drawn in each practical syllogism from the last auxiliary premise (“This applicant is black”) to the conclusion (“I must not hire this applicant”), an inference that consists in taking a person’s race or sex as a reason to treat her less favorably than one would treat others. (“*Because* this applicant is black, *therefore* I must not hire this applicant.”) Antidiscrimination law does not, of course, prohibit the mere drawing of such an inference, as if the law could legitimately and effectively proscribe people’s thoughts. Antidiscrimination law instead forbids people to adopt the inference as their plan of

action and then successfully to execute that plan.³⁵ Nevertheless, while antidiscrimination law prohibits acts rather than mental features, it identifies the acts it prohibits in terms of a mental feature of the people who perform them—namely, the logical inference that one should, on a particular occasion, treat a person less favorably than one would treat others because she is of a particular race or sex.

How should we characterize this inference and the decision to act on it? What kind of mental feature is it, exactly? One possibility is that it is the *motive* of the person who acts on it. According to Steven Sverdlik, “the motive of an action is the agent’s ultimate desire for performing it. Such a psychological state establishes the agent’s end in acting.”³⁶ Motives are, on this definition, subjective mental states rather than objective ones: they reflect what “the agent took to be desirable—and hence, presumably, reasonable—about acting,” not what the agent should have found desirable about acting or what was in fact reasonable for him to do.³⁷ In each of the practical syllogisms considered above, what the agent took to be desirable about acting was embodied in the first, operative premise (or perhaps some still more ultimate, but unexpressed, premise). The agents’ motives thus correspond to the major, operative premises of the practical syllogisms.³⁸ By contrast, the conclusions of the practical syllogisms embody the agents’ “derivative desires,” their desires for the means that help them to realize their ultimate desires in acting.³⁹

Thus defined, motives cannot be the mental features with which antidiscrimination law is concerned. For, as we have seen, the legislation does not directly consider the operative premises of the practical syllogisms on which discriminators act,

³⁵ According to Aristotle, the conclusion of a practical syllogism issues in immediate action (see Aristotle, *Nicomachean Ethics*, trans. Christopher Rowe (Oxford: Oxford University Press, 2002), 1147a27).

³⁶ Steven Sverdlik, *Motive and Rightness* (Oxford: Oxford University Press, 2011), 19.

³⁷ *Ibid.*, 24.

³⁸ *Ibid.*, 27.

³⁹ *Ibid.*

prescinding from both “the emotional or attitudinal force behind what was done” and “the further and ultimate intention with which it was done.”⁴⁰ For purposes of antidiscrimination law, it matters not whether one discriminates to vent one’s animus or prejudice, to maximize productive efficiency, or to placate a bigoted clientele. So long as one decides to treat a member of one race or sex less favorably than one would treat the members of other races or the other sex, one commits direct discrimination and violates antidiscrimination law.

Like most contemporary moral philosophers, Sverdlik distinguishes motives from intentions, so perhaps the latter mental states are the ones with which antidiscrimination law is concerned. According to Sverdlik, whereas a motive is an agent’s ultimate reason for acting, an “intention is . . . a distinct sort of psychological state one of whose functional roles is the initiation of action. In forming such an intention the agent ratifies or endorses a desire in a distinctively practical way by deciding to bring about what she wants to do.”⁴¹ John Finnis explicates the “distinctively practical” attitude that characterizes intentions in terms of the notion of choice. More specifically, he defines an intention as “the linking of means and ends in a plan or *proposal-for-action adopted by choice* in preference to alternative proposals (including to do nothing).”⁴² In other words, “[w]hat one intends is what one chooses, whether as an end or as means. Included in one’s intention is everything which is part of one’s plan (proposal), whether as purpose or as way of effecting one’s purpose(s)—everything which is part of one’s reason for behaving as one does.”⁴³ On this view, intentions differ from motives in at least two respects. First, intentions sweep more broadly than motives, encompassing “not only

⁴⁰ Gardner, “On the Ground,” 181 n.34. Cf. John Finnis, *Intention and Identity*, vol. 2 of *Collected Essays* (Oxford: Oxford University Press, 2011), 274.

⁴¹ Sverdlik, *Motive*, 33-34 (footnote omitted).

⁴² John Finnis, “Intention in Tort Law,” in *Philosophical Foundations of Tort Law*, ed. David G. Owen (Oxford: Oxford University Press, 1995), 229.

⁴³ *Ibid.*

one's ultimate ends"—that is, one's motives—"but also all the means one has chosen to further those ends."⁴⁴ Second, as Sverdlik suggests, intentions are, in a sense, more "practical" than motives: whereas one forms a motive when one recognizes that a particular plan of action will bring about a particular end that one desires, one does not form an intention until one settles on that plan of action, until one chooses it as the option one will pursue.

Although intentions and motives thus differ in some important respects, they are similar in another one. In particular, both intentions and motives reflect what an agent regards as his *reasons* for performing a particular action. This truth is obvious in the case of motives, which simply are what an agent takes to be his ultimate reasons for acting. And because an agent's motives and ultimate intentions have the same "content,"⁴⁵ an agent's ultimate intentions also reflect what he takes to be his ultimate reasons for acting. Indeed, all intentions reflect what the agent regards as his reasons for action, if not always his *ultimate* reasons.⁴⁶ As Elizabeth Anscombe explains, an agent's intention in performing a particular action is the answer he would give if he were asked *why* he was performing that action, an answer that would necessarily take the form of a reason (unless, of course, his answer were simply, "I don't know," in which case his action would be unintentional).⁴⁷ To be sure, an agent could give multiple, mutually consistent answers to this "why" question. For example, if we asked a person why he was getting into his car, he could respond, "To go to the doctor," "To cure my cold," or "To restore my health," and each of these answers could describe his intention with equal accuracy. This multiplicity of available responses does not embarrass an account that equates an agent's intentions with what he regards as his reasons for action. Rather, it simply highlights the

⁴⁴ Ibid., 232.

⁴⁵ Sverdlik, *Motive*, 34.

⁴⁶ See Joseph Raz, *Engaging Reason* (Oxford: Oxford University Press, 2002), 39 ("[T]o identify people's intentions one needs to identify their reasons. They fix the intentions.").

⁴⁷ See G.E.M. Anscombe, *Intention*, 2nd ed. (Oxford: Basil Blackwell, 1963), 24-25.

fact that an agent intends both the ultimate end he seeks to realize through an action and all the means he chooses to realize that end, so that each means stands as an end with respect to another, more immediate means in a chain of means-ends relations culminating in the ultimate end.⁴⁸ We can intelligibly identify any of the means in this chain as part of the agent's intention and thus among what he perceives to be his reasons for acting, even if the agent understands each means as a reason only in relation to the ultimate end he is pursuing—that is, his motive.

Many political and legal theorists use the term “intention” in a second, looser sense to refer not to what an agent takes to be his reasons for action, but to how he regards the physical components and effects of his actions. According to the latter use, an agent intends to perform an action so long as he is aware of what he is doing and appreciates the consequences he is bringing about.⁴⁹ As we saw in Chapter 3, many theorists of tort law espouse this understanding of intentions, supposing that an agent acts with the requisite mental state for intentional torts so long as he foresees (or should foresee) the harm his action will cause. This understanding of intention is also highly congenial to antiperfectionism, since we can determine whether an agent was aware of what he was doing and the consequences he was bringing about without considering his character.

Antidiscrimination law, however, adopts the stricter understanding of intentions rather than the looser one, focusing on at least some of the reasons for which people act. If antidiscrimination law instead adopted the looser understanding of intentions, then it would consider the conclusions of the practical syllogisms outlined above in complete isolation from the other premises, asking only whether the agent was aware of the action he was performing—say, refusing to hire a job applicant—and the consequences that

⁴⁸ See Finnis, *Intention*, 159-63.

⁴⁹ See Scanlon, *Moral Dimensions*, 10.

action entailed. In fact, at least prohibitions against direct discrimination ask *why* the agent performed the action—why, for instance, he refused to hire the job applicant. If the agent could truthfully answer, in part, “Because the applicant was black” or “Because the applicant was a woman,” then he has committed direct discrimination based on race or sex.⁵⁰ Antidiscrimination law, in other words, forbids certain relevantly situated people to take another person’s race or sex as a reason to treat her less favorably than one would treat people of other races or the other sex. It matters not, moreover, whether a discriminator regards another person’s race or sex as a means or end; race or sex need not be his ultimate reason—that is, his motive—but rather merely one of his reasons. In each of the examples above, because the agent regards the applicant’s race as a (derivative or ultimate) reason for not hiring the applicant, the agent acts on a discriminatory intention.

Indirect Discrimination

We might worry that, with its focus on the intentions of discriminators, the preceding account of antidiscrimination law fails to explain prohibitions against indirect (or “disparate impact”) discrimination. For such prohibitions seem largely to prescind from the intentions of discriminators, attending instead to the effects certain rules, criteria, procedures, and the like have for racial minorities and women. It is this very concern that prompts many proponents of systemic theories of discrimination to disparage prohibitions against direct discrimination and to exalt prohibitions against indirect discrimination.⁵¹

To note that prohibitions against direct discrimination focus on the reasons for which discriminators act, however, is not to deprive prohibitions against indirect discrimination of a robust justification. On the contrary, we can integrate prohibitions

⁵⁰ Cf. Raz, *Engaging Reason*, 40 (“To show that an act is intentional it is not enough to show that its initiation and conduct are under the agents’ control. They also must have a story to tell about it, a story which makes its performance intelligible. It explains why they exercised their control to perform it, rather than to avoid it.”).

⁵¹ See, e.g., Freeman’s discussion of indirect discrimination in “Legitimizing,” 1093-99. See also Christopher McCrudden, “Institutional Discrimination,” *Oxford Journal of Legal Studies* 2 (1982): 303-67.

against indirect discrimination into a reason-based account of prohibitions against direct discrimination in a number of ways. For example, if, following the so-called doctrine of double effect, we accept the distinction between intention and foresight, then we can conceive of prohibitions against indirect discrimination as necessary measures for regulating the foreseeable, but unintended (in the strict sense of “intention”), consequences of people’s actions, which prohibitions against direct discrimination altogether ignore.⁵² Even if an agent does not regard another person’s race as one of his reasons for acting, it still might be apparent that one of his actions will burden racial minorities more severely than white people—because, say, racial minorities are significantly more likely to possess some other trait that the agent *does* take as one of his reasons for action. And yet, we still might want to require him to avert the foreseeable disparate impact on racial minorities (unless he can adduce a sufficiently compelling race-neutral justification for not doing so). We can understand prohibitions against indirect discrimination as performing this function.

According to this account, prohibitions against indirect discrimination do not regulate the reasons for which people act. Some theorists, however, might wish to argue that prohibitions against indirect discrimination perform precisely this function—that just like prohibitions against direct discrimination, they forbid people to act for certain reasons, if not the same reasons that make direct discrimination wrongful. More specifically, we can understand prohibitions against indirect discrimination as expanding the set of forbidden grounds for differential treatment to encompass not just the standard forbidden grounds of race and sex, but also intrinsically permissible ones that, when made the basis of decision, tend to have a disproportionate impact on racial minorities and

⁵² See Finnis, *Intention*, 274.

women.⁵³ The prohibitions would, for instance, forbid an employer to regard an applicant's education level as a reason for not hiring the applicant if that trait correlated strongly with race (at least without a compelling, race-neutral justification). On this view, we can explicate the wrongfulness of both direct and indirect discrimination in terms of the intentions of discriminators, the reasons for which they act.

Note, however, that even a reason-based account of indirect discrimination diverges from the account of direct discrimination offered above in an important respect: whereas prohibitions against direct discrimination regard certain intentions as intrinsically wrong, irrespective of any consequences acting on them might have, prohibitions against indirect discrimination regard certain intentions as, at most, contingently wrong, wrong only insofar as acting on them produces certain undesirable consequences. Thus, even if we adopt a reason-based account of prohibitions against indirect discrimination, we shall have to ground them in a different justification from the one we would adduce for prohibitions against direct discrimination. In particular, we shall have to explain why it is wrong to act on certain race- and sex-neutral intentions only when doing so disproportionately burdens racial minorities or women.

The most plausible explanation is that the disproportionate burdens constitute a distributive injustice. On this view, prohibitions against indirect discrimination forbid people to act for certain race- and sex-neutral reasons in certain circumstances in order to prevent them from creating or contributing to distributive injustice—or, conversely, in order to compel them to help to promote distributive justice. Prohibitions against indirect discrimination will serve the same purpose even if we construe them in terms of foresight

⁵³ See, e.g., Elisa Holmes, "Anti-Discrimination Rights Without Equality," *Modern Law Review* 68 (2005): 190-91.

rather than intention.⁵⁴ Either way, prohibitions against direct discrimination and prohibitions against indirect discrimination serve different functions and thus will require different justifications.

If prohibitions against indirect discrimination are indeed to be justified according to principles of distributive justice, then antiperfectionism should have no difficulty legitimating them. For contemporary liberal theories of distributive justice *assess* only the consequences of actions, not people's reasons for acting. To be sure, a theory of distributive justice will inevitably have to *identify* people's intentions in order to determine their causal significance, their tendency, when acted on, either to foster or to impede distributive justice.⁵⁵ But such identification need not involve assessment and thus presents no threat of perfectionism. Because prohibitions against indirect discrimination seem so obviously to accord with antiperfectionism, the rest of this chapter will focus on direct discrimination, which requires us to assess the reasons for which discriminators act and thus potentially poses a greater challenge for antiperfectionism.

Some antiperfectionists might seize on this concession and attempt to cabin prohibitions against direct discrimination into the structure of prohibitions against indirect discrimination, so that the former come to share the latter's distributive concerns. But as we have seen, the provisions of antidiscrimination law themselves define direct discrimination in terms of the intentions of discriminators rather than the consequences—distributive or otherwise—of discriminatory acts. If we are to justify prohibitions against direct discrimination, we shall have to move beyond the distributive focus of prohibitions against indirect discrimination and explain why taking another person's race or sex as a

⁵⁴ Most theorists seem to agree that prohibitions against indirect discrimination perform a redistributive function, though they disagree about whom, exactly, goods are being redistributed among. See John Gardner, "Liberals and Unlawful Discrimination," *Oxford Journal of Legal Studies* 9 (1989): 8-11.

⁵⁵ Cf. Scanlon's notion of the "predictive" significance of intentions (*Moral Dimensions*, 13).

reason to treat her less favorably than one would treat people of other races or the other sex is intrinsically wrong.

V. INTENTION AND MORAL PERMISSIBILITY

At least through its prohibitions against direct discrimination, antidiscrimination law forbids people to act for certain reasons. For example, whereas an employer may ordinarily refuse to hire a person for any reason, antidiscrimination law forbids him to do so for the particular reason that the person happens to be black or a woman. This kind of prohibition assumes that the very same act—say, refusing to hire someone—can be either right or wrong, depending on the reasons for which the agent performs it—that what an agent takes to be his reasons for performing a particular act partly determines the act’s permissibility.

And yet, some philosophers deny that one can transmute an ordinarily permissible action into an impermissible one simply by performing it for some illicit reason. Judith Thomson, for instance, espouses what she dubs “The Irrelevance-of-Intention-to-Permissibility Thesis,” according to which “[i]t is irrelevant to the question whether X [an agent] may do alpha [an action] what intention X would do alpha with if he or she did it.”⁵⁶ Somewhat more weakly, T.M. Scanlon denies that an action’s permissibility can depend “fundamentally” on the intentions of the agent who performs it,⁵⁷ though he concedes the existence of certain “derivative” connections between moral permissibility and intentions.⁵⁸ Despite this difference between Thomson’s position and Scanlon’s, however, both threaten to preclude a perfectionist account of discrimination, which

⁵⁶ Judith Jarvis Thomson, “Self-Defense,” *Philosophy & Public Affairs* 20 (1991): 295.

⁵⁷ Scanlon, *Moral Dimensions*, ch. 1.

⁵⁸ For example, the intention with which an agent performs an action might have “predictive” significance, in that it might indicate the likely consequences of the action, which are undoubtedly relevant to the action’s permissibility. An agent’s intention might also determine the “meaning” of his action—the significance of the action for his relations with others—which, when mediated through other moral principles, can be relevant to the action’s permissibility. See *ibid.*, ch. 2.

supposes that actions can indeed be wrong simply in virtue of the reasons for which people perform them.

Thomson and Scanlon do not claim that people's intentions are irrelevant to all forms of moral assessment. Rather, they seek to distinguish the permissibility of acts from other dimensions of moral assessment and to consign the assessment of intentions to those other dimensions. For example, Thomson allows that a person's intentions are relevant to determining whether he is at fault for performing a particular action, but she insists that "[i]t is irrelevant to the question whether X may do alpha whether X would be at fault in doing it."⁵⁹ To condemn a person's character, to brand him a bad person, is not to forbid his action, whose permissibility in no way depends on such considerations.⁶⁰

Rather than simply distinguish between assessments of acts and assessments of persons, Scanlon draws a narrower distinction between two uses of moral principles. In particular, we might use moral principles either "as standards of criticism or as guides to deliberation."⁶¹ According to Scanlon, when we use moral principles critically, in order to assess the process by which a person went about deciding what to do, we must indeed consider what the person took to be his reasons for action.⁶² But when we use moral principles deliberatively, in order to decide what a person should do or should have done in certain circumstances, we need attend only to the "considerations" that the principles identify as "count[ing] for or against various courses of action," "considerations" that include the (moral and physical) effects of an action but not the process by which an agent went about deciding whether to perform it.⁶³ It is this second, deliberative kind of assessment, Scanlon contends, that determines the moral permissibility of actions.⁶⁴

⁵⁹ Thomson, "Self-Defense," 295.

⁶⁰ Judith Jarvis Thomson, "Physician-Assisted Suicide: Two Arguments," *Ethics* 109 (1999): 516-17.

⁶¹ Scanlon, *Moral Dimensions*, 22.

⁶² *Ibid.*, 23.

⁶³ *Ibid.*

⁶⁴ *Ibid.*, 23, 27.

Delineating the moral terrain along lines highly congenial to antiperfectionism, both Thomson and Scanlon decouple assessments of moral permissibility from assessments of moral character and consign questions of intention exclusively to the latter dimension.

Why do philosophers like Thomson and Scanlon demand that we assess moral permissibility separately from moral character? Why, in particular, do they resist predicating the moral permissibility of actions on the intentions with which they are performed? Thomson, for one, offers little argument for her position, instead appealing to our intuitions about various thought experiments. Nevertheless, she does claim that the contrary position, according to which an illicit intention can corrupt an otherwise permissible action, has at least one “very odd” implication: it appears to require agents to focus not on the features of the actions they perform, but rather on the reasons for which they perform those actions.⁶⁵ Once again, Thomson does not explain precisely why she considers this implication “very odd,” but perhaps she thinks it imposes too onerous a requirement on moral agents, that morality would be overly demanding were it to insist not merely that people perform only ordinarily permissible actions, but also that they continually scrutinize their psyches to ensure that they do so only for good reasons.⁶⁶

This conclusion, however, is too quick. For while it might indeed be overly demanding (if not completely hopeless) to expect people to conjure good reasons for performing an action when they otherwise perceive none, it seems perfectly reasonable to require that they recognize when they would be acting for a bad reason and refrain from doing so. Therefore, to dismiss as draconian any moral theory that regulates our intentions as well as our actions is arbitrarily to preclude moral requirements we are perfectly capable of fulfilling.

⁶⁵ Thomson, “Self-Defense,” 293.

⁶⁶ Cf. Scanlon, *Moral Permissibility*, 58-62.

Whatever their motivations for bifurcating assessments of moral permissibility and moral character, both Thomson and Scanlon ignore certain actions whose impermissibility we cannot explain without reference to the reasons for which they are performed. Consider the simple act of touching a stranger's shoulder. One could perform this act for any number of reasons, intending it as "a greeting, a warning, an encouragement, a condescension," and so on.⁶⁷ The permissibility of the act depends on which of these reasons motivated the agent. We would, for example, generally consider it permissible for one to touch a stranger's shoulder as a warning but not as a means of sexual arousal. And yet, we cannot distinguish between these two instances of touching, let alone assess them, without attending to the agent's intentions, the reasons for which he acted.

Or consider one of Scanlon's examples.⁶⁸ Suppose a person promises to sell his house to his friend but then breaks his promise after receiving a better offer from a stranger. Suppose further that we are generally required to keep our promises but that this requirement can be defeated by legitimate "countervailing considerations." So long as no such considerations were present when the seller broke his promise to his friend, Scanlon argues, his action was impermissible—not because the seller broke his promise for a selfish reason (viz., getting a better price for his house), but because he broke his promise in the absence of legitimate countervailing considerations. While the seller's selfish intention may indeed be relevant to a *critical* assessment of the seller's action, an assessment of the process by which the seller went about deciding whether to break his promise, it is, Scanlon contends, irrelevant to a *deliberative* assessment, an assessment of the considerations that applied in the situation and that determined the action's

⁶⁷ Finnis, "Intention in Tort Law," 239.

⁶⁸ Scanlon, *Moral Dimensions*, 23-24.

permissibility.⁶⁹ But suppose that some legitimate countervailing consideration was present when the seller broke his promise. Although the seller could have broken his promise for either this consideration or the selfish reason of getting a better price, we believe that breaking his promise would be permissible only if he acted for the former reason. The permissibility of the seller's action would depend on the quality of the reason for which he acted.⁷⁰

Our intentions, the reasons for which we act, can determine the permissibility of at least some of our actions. Thomson and Scanlon's contrary conclusion rests on "the common unsound assumption that human acts can be identified, for purposes of moral or legal assessment, independently of the acting person's intention(s). To be sure, outward *behavior* can be so identified. But countless *acts* cannot be truly identified for what they are (*prior* to assessment as right or wrong, lawful or unlawful) unless and until the outward behavior which they involve is understood as *the carrying out of such and such an intention.*"⁷¹ If we cannot even properly identify at least some acts without considering the intentions of the people who perform them, then, a fortiori, we cannot determine their permissibility without considering those intentions.

VI. DISCRIMINATION AND INTENTION

Even if there is a category of actions whose permissibility depends on the reasons for which people perform them, we might nonetheless wonder whether the kinds of actions proscribed by antidiscrimination law fall into that category. Does the permissibility of, say, an employer's refusing to hire a particular job applicant really depend on whether the employer takes the applicant's race or sex as a reason for the

⁶⁹ *Ibid.*, 24.

⁷⁰ Scanlon would respond that the mere presence of the countervailing consideration vitiates the seller's obligation to keep his promise, even if he ultimately breaks the promise for some other, illegitimate reason (cf. *ibid.*, 65-66). This assertion is puzzling. It seems more intuitive to say that the presence of the countervailing consideration allows the seller to break the promise only for that reason, not just for any reason.

⁷¹ Finnis, "Intention in Tort Law," 239.

refusal? To answer this question, we must address two distinct issues: (1) whether we can identify instances of discrimination without considering the intentions of discriminators and (2) whether we can explain the wrongfulness of discrimination without considering those intentions. This section will argue that we must consider discriminators' intentions, the reasons for which they act, in order both to identify and to condemn discrimination.

Intention and the Identification of Discrimination

Deborah Hellman defends the opposite conclusion, seeking to purge from accounts of discrimination all consideration of people's intentions. Hellman first argues that discriminators' intentions are irrelevant to identifying instances of discrimination because "an actor's intention does not determine whether he or she in fact distinguishes on the basis of a particular trait," a necessary prerequisite for discrimination.⁷² On its face, this assertion is puzzling, if not incoherent. For "distinguish," like "discriminate," is an intentional verb: one "distinguishes on the basis of a particular trait" when one perceives the presence of that trait in someone and takes it as a reason to contrast him with other people who lack it. If we ignore people's intentions altogether, how can we determine whether they have actually distinguished among others based on race or sex and committed discrimination?

Consider the following example, devised by Larry Alexander. Attempting to dissemble its discriminatory intentions, a city excludes black people from its public pool not by promulgating an explicitly discriminatory access rule, but rather by instructing the pool's lifeguard to open the pool whenever white people want to swim there and to close it immediately once any black people arrive. According to Alexander, because neither the act of opening the pool nor the act of closing it, taken by itself, is discriminatory, we can determine whether the exclusion of black people from the pool constitutes discrimination

⁷² Hellman, *When Is Discrimination Wrong?*, 139.

only by identifying the “real” or “operative” access rule—namely, the city’s “metarule” instructing the lifeguard “to switch back and forth between enacted rules.” And we can identify this “metarule,” Alexander contends, only by considering the city’s intentions—its taking black people’s race as a reason to exclude them from the pool (whatever its further motives for doing so).⁷³

Responding directly to Alexander’s pool example, Hellman insists that we can identify every instance of discrimination without scrutinizing the intentions of its perpetrators. Indeed, she argues that one can discriminate unintentionally, not taking a person’s race or sex as a reason for action yet treating that person differently from how one would treat people of other races or the other sex. For, according to Hellman, discrimination consists not in intentional disparate treatment, but rather in disparate treatment simpliciter—not in what reasons people perceive and act on, but in “what takes place.”⁷⁴ On this view, a decision in which race or sex figures only subconsciously, without the decisionmaker’s perceiving its influence, is still a discriminatory one.⁷⁵ So, too, is a decision that is completely unadulterated by race or sex but that, because of some contingency, nevertheless distinguishes between black people and white people or women and men.⁷⁶

Hellman is drawn to this conclusion partly by her eagerness to condemn unconscious forms of discrimination. Were we to define discrimination in terms of the intentions of its perpetrators, Hellman worries, we would have to discount unconscious instances of the conduct, instances where race or sex distorts one’s treatment of others without one’s realizing it.⁷⁷ But is Hellman’s anxiety justified? Can we not explicate even

⁷³ Larry Alexander, “Rules, Rights, Options, and Time,” *Legal Theory* 6 (2000): 399-400.

⁷⁴ Hellman, *When Is Discrimination Wrong?*, 145.

⁷⁵ *Ibid.*, 141-43.

⁷⁶ *Ibid.*, 145.

⁷⁷ For a recent overview of the problem of unconscious discrimination—particularly the problem of “implicit bias”—and the relevant psychological literature, see Jerry Kang and Kristin Lane, “Seeing

unconscious discrimination as at least sometimes stemming from the intentions of its perpetrators? Hellman reaches the opposite conclusion only because she conflates two different distinctions: the intentional/unintentional distinction and the conscious/unconscious distinction. In particular, she conceives of intentions as conscious mental states, contrasting them with unconscious cognitive propensities.⁷⁸ If intentions are conscious by definition, if an agent can never form an intention implicitly, then it does seem that we must either condone unconscious discrimination or redefine all forms of discrimination independently of their perpetrators' intentions.

To insist that our intentions can never be implicit, however, is to misapprehend their nature. Recall that the account of discrimination expounded in Section IV above explicated each instance of the conduct in terms of its perpetrator's reasoning, which, in turn, was depicted as a practical syllogism comprising a race- or sex-based operative premise; a purely informational, auxiliary premise; and a practical conclusion that bore the operative premise's race- or sex-based taint. While we might assume that this reasoning process must manifest itself explicitly in the perpetrator's mind, such an assumption too literally construes "the model of the practical syllogism," which "is very often an ideal reconstruction rather than a conscious process."⁷⁹ The perpetrator need not consciously perceive his intention, the logical inference from the auxiliary premise identifying a person's race or sex to the conclusion that he should treat her differently from how he would treat people of other races or the other sex. Intentions do not always function as "mental causes," the conscious cognitive events that actually prompt people to

Through Color Blindness: Implicit Bias and the Law," *UCLA Law Review* 58 (2010): 465-520. For the classic statement of the problem, see Charles R. Lawrence III, "The Id, the Ego, and Equal Protection: Reckoning with Unconscious Racism," *Stanford Law Review* 39 (1987): 317-88.

⁷⁸ See, e.g., Hellman, *When Is Discrimination Wrong?*, 139-40.

⁷⁹ Steven Sverdlik, "Motive and Rightness," *Ethics* 106 (1996): 338.

act.⁸⁰ Sometimes, our actions are motivated by our emotions—or even factors external to us—rather than by what we regard as our reasons for action.

And yet, we do not automatically dismiss such actions as unintentional, as disconnected from what we regard as our reasons for action. On the contrary, as Philip Pettit rightly notes, we suppose that people exercise a kind of unconscious, “virtual” control over many of their actions, a control that enables them to act intentionally even when they do not explicitly ponder the precise reasons for their conduct. That is, in many cases where they are “being ruled by habit and routine, achieving their goals with minimal attention and effort,” people nonetheless act for reasons; they still adopt plans of action comprising ends (“goals”) and means.⁸¹ For were their respective habits or routines somehow to be upended, were they suddenly to discover themselves performing actions they had not at least implicitly endeavored to perform, they would often be provoked to reflect consciously on their intentions and either to acquiesce in their new, altered patterns of behavior or to revert to their original habits or routines.⁸² Or as Raz puts it, it is not the case that

we are active only when we are self-reflective and explicitly hold views about our rationality, let alone that we are active only when we have views about the nature of rationality. We believe that we are properly responsive to reason whenever we would, if challenged, resist, at least initially, claims that our conduct, intentions, emotions, etc., are irrational or unreasonable.⁸³

For example, while an employer might discriminate only habitually, unconsciously treating a black person differently from how he would treat a white person, he nonetheless acts intentionally so long as he would be rationally committed to regarding

⁸⁰ Anscombe, *Intention*, 17-18.

⁸¹ Philip Pettit, *A Theory of Freedom: From the Psychology to the Politics of Agency* (Oxford: Oxford University Press, 2001), 92.

⁸² See *ibid.*, 38-39, 76, 91-93. Cf. Gardner’s argument that one can be fully justified even if one acts “by instinct or habit” (“Justifications and Reasons,” in *Harm and Culpability*, eds. A.P. Simester and A.T.H. Smith (Oxford: Oxford University Press, 1996), 112 n.20). For another argument that one can be responsive to reasons without being conscious of them, see George Sher, *Who Knew? Responsibility Without Awareness* (Oxford: Oxford University Press, 2009), 66-69.

⁸³ Raz, *Engaging Reason*, 17.

the black person's race as a reason to treat that person less favorably than one would treat people of other races when prompted to reflect on his conduct, so that his race-based differential treatment was either an end in itself or a means to some ulterior end. Hellman's argument to the contrary notwithstanding, at least some unconscious discrimination can ultimately owe to its perpetrators' intentions.

Hellman wishes her definition of discrimination to encompass not only unconscious discrimination, but also any action that has a discriminatory "effect." Recall that Hellman argues that the access rule in Alexander's pool example is discriminatory so long as it has the effect of excluding black people—irrespective of its author's intentions. Suppose, for instance, that the lifeguard consistently closes the pool when black people arrive neither because his employer instructed him to do so nor because he himself suffers from (conscious or unconscious) racial bias, but rather because it just so happens that (unknown to the lifeguard) the local black population has time to use the pool only at the exact moment when the lifeguard regularly takes his lunch break. On Hellman's account, although the lifeguard took the time, not anyone's race, as his reason for action, he still discriminates based on race. We (and the law) might be inclined to consider his conduct to be an instance of indirect, rather than direct, discrimination, but Hellman seems to conflate the two, defining discrimination to include any conduct that has a discriminatory effect.⁸⁴

To avoid this conflation while nevertheless prescind from the intentions of discriminators, Hellman might argue that an effect is discriminatory whenever it was *caused* by race or sex—that is, whenever the effect would not have materialized "but for"

⁸⁴ At some points, the conflation is nearly explicit. See Hellman, *When Is Discrimination Wrong?*, 145-46, 146 n.22. It does not help matters that Hellman misstates the legal distinction between direct and indirect discrimination. For example, she defines cases of direct (or "disparate treatment") discrimination as "those in which the law, policy, or action *explicitly* draws a distinction on the basis of a characteristic legally recognized as suspect" (*ibid.*, 147 (emphasis added)). In fact, according to U.K. and U.S. antidiscrimination law, even non-explicit, covert discrimination can still be direct discrimination. The discrimination might have to be indirectly *proved*, but it remains direct discrimination.

race or sex. Andrew Koppelman analyzes discrimination in similar terms. Attempting to equate discrimination based on sexual orientation with discrimination based on sex, Koppelman argues that one is a victim of sex discrimination whenever one's sex is a necessary condition for the differential treatment one suffers, whenever one would not have suffered the differential treatment had one been of the other sex.⁸⁵ In other words, one is a victim of sex discrimination whenever one's sex was a "but for" cause of one's differential treatment. Prescinding from the intentions of discriminators, this account apparently renders intuitively correct verdicts in particular cases, deeming pregnancy discrimination, for instance, a form of sex discrimination. For although an employer might insist that he is taking pregnancy, not sex, as his reason for treating his pregnant employees less favorably, they nonetheless would not have suffered the differential treatment had they not been women.

This account of discrimination, however, still subsumes under the heading of "direct discrimination" some forms of conduct that the law treats as indirect discrimination. For one's race or sex can be a "but for" cause of one's differential treatment even if one's race or sex was not taken as a reason to treat one less favorably—for instance, when an employer or purveyor of a public accommodation bases his decision on an ostensibly race- or sex-neutral legal category like pensionable age, which, in turn, has been defined in terms of race or sex.⁸⁶ And yet, most contemporary western liberal democracies treat direct and indirect discrimination as two distinct kinds of wrongs. To be sure, this objection by no means refutes a causal account of discrimination; extant antidiscrimination law could simply be misconceived. But we do tend to think that direct

⁸⁵ Andrew Koppelman, *Antidiscrimination Law and Social Equality* (New Haven: Yale University Press, 1996), 154-58. For a similar argument equating discrimination based on sexual orientation with discrimination based on sex, see Robert Wintemute, *Sexual Orientation and Human Rights* (Oxford: Oxford University Press, 1995), 202-03. The causal account of discrimination has also found some support in the law. In U.K. law, see *James v. Eastleigh Borough Council*, [1990] 2 AC 751.

⁸⁶ This was the case in *Eastleigh Borough Council*.

and indirect discrimination are two distinct kinds of wrongs, a judgment that a causal account of discrimination calls into question.⁸⁷

The preceding objection suggests that the causal definition of discrimination is overinclusive, that it subsumes under a single heading genuinely distinct forms of conduct. The causal definition is also underinclusive, failing to account for kinds of conduct that our intuitions and the law regard as discrimination. Most notably, the causal definition of discrimination cannot explain cases of so-called mixed-motives discrimination,⁸⁸ cases where a discriminator has multiple reasons for treating a person less favorably than others, each of which is sufficient for his decision and, therefore, neither of which is necessary.⁸⁹ An employer might, for instance, fire one of his employees both because the employee is a woman and because she is rude; he might, in other words, take both her sex and rudeness as reasons for action. It might be the case, moreover, that the employer would have made the very same decision had the employee still been rude but not a woman. Because sex was not a “but for” cause of the employee’s termination, because she would have been terminated even had she been a man, a causal account will not consider the employer’s conduct to be an instance of discrimination. Our intuitions and the law hold otherwise. Thus, to identify all the instances of discrimination recognized by our intuitions and the law, we must define discrimination in terms of people’s intentions, the reasons for which they act.

Intention and the Morality of Discrimination

⁸⁷ For other versions of this objection, see Cavanagh, *Against Equality of Opportunity*, 199; and Iris Young, *Justice and the Politics of Difference* (Princeton: Princeton University Press, 1990), 196. Not everyone shares this belief. Sophia Moreau, for instance, holds what she considers to be the “very fundamental intuition . . . that both forms of discrimination are forms of the same thing, the same kind of injustice” (“What Is Discrimination?,” 176).

⁸⁸ Despite the law’s use of the term “motives,” the better term would be “intentions,” since the law considers only whether race or sex was *one* of the discriminator’s reasons, not whether it was his *ultimate* reason (i.e., his motive).

⁸⁹ In U.S. law, see § 703(m) of Title VII of the Civil Rights Act of 1964, as amended by the Civil Rights Act of 1991. Race or sex need not have been a “but for” cause for a mixed-motives claim to lie. See *Desert Palace, Inc. v. Costa*, 539 U.S. 90 (2003).

Even if we must define discrimination in terms of the intentions of its perpetrators, it does not necessarily follow that we must locate the conduct's wrongfulness there as well. In other words, while certain conduct might stem from discriminatory intentions, it might be wrongful in virtue of its other features. Following Thomson, Scanlon, and other likeminded theorists, Hellman espouses just such a position, insisting that the immorality of a discriminatory act inheres not in its perpetrator's intentions, but rather in "the fact that the race of [the victim] affects a decision when it should not."⁹⁰ More specifically, Hellman contends that a person cannot transmute an otherwise permissible action into an impermissible one simply by intending it as a means of discriminating based on race or sex.⁹¹ To be sure, in performing a permissible action with such an intention, the person might evince a defective moral character, an unwarranted contempt, say, toward some of his fellow human beings.⁹² But, Hellman asserts, to exhibit fault is not necessarily to act impermissibly; at least in cases of discrimination, assessments of moral permissibility and assessments of moral character often diverge.

As comfortingly nonmoralistic as this position initially seems, it actually renders counterintuitive verdicts in many particular cases, condoning forms of discrimination that antidiscrimination law and our ordinary moral intuitions condemn. In particular, because many furtive forms of discrimination outwardly resemble perfectly innocent kinds of conduct, we can morally distinguish the former from the latter only by scrutinizing the intentions of their respective perpetrators. Suppose, for example, that the owners of two different businesses both refuse to hire applicants who reside outside the state in which the businesses are located but that they do so for very different reasons. Noting the superior quality of the state's school system, the first employer uses state residency as a proxy for educational attainment, a proxy that, if empirically sound, is surely morally

⁹⁰ Hellman, *When Is Discrimination Wrong?*, 154.

⁹¹ See, e.g., *ibid.*, 166.

⁹² See, e.g., *ibid.*, 159, 166.

permissible. The second employer, by contrast, cares more about the whiteness of his workforce than its competence and hires only residents of the state simply because so few of them are black.⁹³ Were we to follow Hellman in prescinding from all moral evaluation of people's intentions, it seems we would have to countenance the second employer's use of the proxy, just as we would the first employer's. For the two employers differ not in terms of the proxy they use, but rather in terms of the reasons for which they use it. We can distinguish between these two uses of the proxy only by examining the employers' respective intentions, the reasons on which each of them acted.

Lest we worry that this criticism construes Hellman's argument uncharitably, Hellman herself explicitly concedes that people may discriminate against black people by proxy so long as "race is not used explicitly" and the proxy "is not understood in a racialized manner."⁹⁴ On this view, an employer need only fabricate some plausible, ostensibly race-neutral pretext to legitimate his efforts to exclude black people from his business. Unless we are willing to accept such a counterintuitive result, unless we are willing to condition the morality of various discriminatory acts on the success of the subterfuges their perpetrators concoct, we should ground our account of the wrongfulness of discrimination in an evaluation of the reasons for which its perpetrators act.

Hellman might respond to this line of criticism by attempting to provide an alternative explanation for our intuitions about the preceding example, an explanation that condemns the second employer's actions without assessing his discriminatory intentions. She might argue, for instance, that the second employer acts impermissibly not because he intends to discriminate based on race, but because he applies a hiring criterion utterly

⁹³ Hellman considers an analogous example involving law-school admissions procedures (see *ibid.*, 157-62), which was devised by Larry Alexander and Kevin Cole in "Discrimination by Proxy," *Constitutional Commentary* 14 (1997): 453-63. But surprisingly, Hellman never responds directly to Alexander and Cole's example; instead, she alters it to make it more congenial to her position. See Hellman, *When Is Discrimination Wrong?*, 158.

⁹⁴ Hellman, *When Is Discrimination Wrong?*, 160.

immaterial to his business purposes, thereby violating a principle of merit or a positional duty unique to employers.⁹⁵

This account of the example, however, is unconvincing. First, it is far from obvious that it is morally impermissible (as opposed to merely inefficient) for an employer to violate a principle of merit in hiring employees. Why should we consider an employer to be under a moral duty to hire the most meritorious applicant rather than, say, a friend who happens to be less qualified?⁹⁶ Second, and more fundamentally, the preceding account of the example equates the second employer's discrimination with other forms of conduct, such as nepotism, that also contravene principles of merit but do not elicit serious moral condemnation.⁹⁷ We tend to think that discrimination is much more seriously wrongful than nepotism and other violations of the merit principle. Indeed, we tend to think that it is the offensiveness of employers' discriminatory intentions, not merely the impertinence of their hiring criteria, that morally distinguishes their conduct from permissible hiring decisions. Hellman's alternative explication of the example cannot account for these intuitions.

If the second employer does in fact commit wrongful discrimination, then it once again appears that assessments of moral permissibility supervene on assessments of people's intentions. For it is his taking people's race as a reason for action alone that distinguishes his conduct from the first employer's. Only by considering people's intentions, the reasons for which they act, can we distinguish wrongful discrimination from other, permissible (or differently wrongful) conduct.

VII. DISCRIMINATION AND RESPECT

Thus far, this chapter has attempted to justify antidiscrimination law's focus on the intentions of discriminators. Lest we obviate the distinctive function

⁹⁵ Cf. Alan Goldman, *Justice and Reverse Discrimination* (Princeton: Princeton University Press, 1979), 34.

⁹⁶ Cf. Cavanagh, *Against Equality of Opportunity*, 20.

⁹⁷ For versions of this argument, see *ibid.*, 24-25; and Scanlon, *Moral Dimensions*, 70-71.

antidiscrimination law purports to perform, subsuming the legislation under a more generic program of social justice, we should formulate our account of discrimination in terms of the perpetrators rather than the victims of the conduct. In particular, we should locate the wrongfulness of discrimination in the intentions of discriminators, the reasons for which they act. For we can neither define discrimination nor distinguish it from other, permissible forms of behavior without attending to the intentions of its perpetrators. Unless we are willing to adopt this focus, we shall fail to condemn many instances of discrimination and to legitimize central facets of contemporary antidiscrimination regimes.

While this argument locates the wrongfulness of discrimination in the intentions of discriminators, it does not identify precisely which feature of discriminatory intentions renders them so objectionable, precisely which moral norms discriminators contravene in taking another person's race or sex as a reason to treat her less favorably than one would treat people of other races or the other sex. We can begin to fill this lacuna by considering the evaluative attitudes (or lack thereof) entailed by discriminatory intentions. In particular, it seems that one cannot take another person's race or sex as a reason to treat her less favorably than one would treat others unless one lacks an appropriate regard for her personhood, her very status as a person. Discriminators, in other words, *fail to respect*—and therefore *disrespect*—their victims. Borrowing Joseph Raz's account of respect for persons, this section will develop a theory of discrimination as disrespect, showing how acting on discriminatory intentions entails not necessarily the presence of malicious feelings or false beliefs, but the absence of respectful attitudes, attitudes in which appropriate respect for persons partially consists.

Respect for Persons

Almost all contemporary accounts of respect for persons begin with the work of Immanuel Kant. In his *Groundwork for the Metaphysics of Morals*, Kant argues that “rational beings are called *persons* because their nature already marks them out as an end in itself, that is, as something that may not be used merely as a means, and hence so far limits all choice (and is an object of respect).”⁹⁸ This principle, one of Kant’s formulations of the “Categorical Imperative,” appears to exhibit all the “features that personal respect is taken to have: that it binds all of us, that [it] is owed to everyone irrespective of their virtues or merits, and that it is especially exigent.”⁹⁹ It is from the notion that persons should always be treated as ends in themselves that many contemporary philosophers attempt to derive duties of respect for persons—indeed, even entire respect-based moralities.¹⁰⁰

As we can see from Kant’s principle of respect for persons, such respect is a kind of “recognition respect” rather than “appraisal respect.” According to Stephen Darwall, appraisal respect is “an attitude of positive appraisal of [a] person either as a person or as engaged in some particular pursuit.”¹⁰¹ A person merits such respect insofar as “he has manifested characteristics which make him deserving of such positive appraisal.”¹⁰² By contrast, to have recognition respect for something is to have “a disposition to weigh appropriately in one’s deliberations some feature of the thing in question and to act accordingly.”¹⁰³ Respect for persons is a kind of recognition respect, since we suppose that persons merit respect simply in virtue of their personhood, irrespective of their other traits. More specifically, we respect persons as persons when we “take seriously and

⁹⁸ Immanuel Kant, *Groundwork of the Metaphysics of Morals*, trans. and ed. Mary Gregor (Cambridge: Cambridge University Press, 1997), AK 4:428.

⁹⁹ Leslie Green, “Two Worries About Respect for Persons,” *Ethics* 120 (2010): 214.

¹⁰⁰ See, e.g., the works of Thomas E. Hill, Jr., especially *Dignity and Practical Reason in Kant’s Moral Theory* (Ithaca: Cornell University Press, 1992); and *Respect, Pluralism, and Justice* (Oxford: Oxford University Press, 2000).

¹⁰¹ Stephen L. Darwall, “Two Kinds of Respect,” *Ethics* 88 (1977): 38.

¹⁰² *Ibid.*, 38-39.

¹⁰³ *Ibid.*, 38.

weigh appropriately the fact that they are persons in deliberating about what to do.”¹⁰⁴

When we fail to “take seriously and weigh appropriately” other people’s personhood in our deliberations but act anyway, we commit not just a rational miscalculation, but a moral wrong. Respect for persons is thus a form of *moral* recognition respect: other people’s personhood places moral restrictions on our actions.¹⁰⁵

Philosophers disagree about precisely which actions duties of respect require and preclude. According to Harry Frankfurt, one can treat another person with respect only if one “deal[s] with him exclusively on the basis of those aspects of his particular character or circumstances that are actually relevant to the issue at hand”—only, that is, if one observes the requirements of “impartiality” and “avoid[s] . . . arbitrariness.”¹⁰⁶ In treating a person arbitrarily by discriminating against him based on some irrelevant trait, Frankfurt believes, one fails fully to respect his personhood.

At least as initially articulated, Frankfurt’s account of disrespectful treatment fails to demarcate a morally problematic class of conduct. For, as Paula Casal notes, Frankfurt’s notion of “relevance” is neither necessary nor sufficient to distinguish wrongful discrimination from certain forms of permissible behavior. It is not necessary because factors like race and sex *are* “relevant to the issue at hand” in many cases of wrongful discrimination—when, for instance, a restaurateur will alienate his racist clientele unless he discriminates against black people. And it is not sufficient because one can make decisions based on at least some of a person’s irrelevant characteristics—say, his taste in music when he is applying for a non-music-related job—without contravening a serious moral requirement.¹⁰⁷ Though sometimes inefficient or irritating, decision making based on irrelevant factors is not necessarily disrespectful.

¹⁰⁴ Ibid.

¹⁰⁵ Ibid., 40.

¹⁰⁶ Harry G. Frankfurt, *Necessity, Volition, and Love* (Cambridge: Cambridge University Press, 1999), 150.

¹⁰⁷ Paula Casal, “Why Sufficiency Is Not Enough,” *Ethics* 117 (2007): 302-03.

More generally, we tend to think that one commits the particular moral wrong of disrespect not whenever one behaves irrationally toward other people, but only when one disregards a particular subset of reasons governing interpersonal conduct. In fact, Frankfurt himself adopts this position, denying that “irrationality as such is immoral.”¹⁰⁸ And yet, he seems to reduce respect to at least one facet of mere rationality: “treating relevantly similar cases alike and relevantly unlike cases differently.”¹⁰⁹ If respecting others consisted in nothing more than heeding this precept, if one could respect others simply by eschewing a particular kind of irrationality, then duties of respect would not constitute a distinct set of moral requirements. Instead, treating another person with respect would be, as Raz puts it, a mere “byproduct” of following “right reason,” of simply doing what reason requires of us anyway.¹¹⁰ Only if duties of respect do not simply reduce to our other moral duties, or to the requirements of rationality more generally, will they have the unique status we ascribe to them.¹¹¹

To distinguish duties of respect from the rest of morality, we must better understand the basis of those duties—namely, our personhood, that which makes persons, as Kant puts it, “ends in themselves.” Raz argues that we should understand this latter notion in terms of the idea of one’s being “valuable in oneself,” one’s “having intrinsic features which endow one with unconditional value, where that value is independent of being good for something or someone else, and its possession is a complete reason for treating its possessor with respect.”¹¹² Raz begins to unpack this account by elaborating the distinction between conditionally and unconditionally valuable goods. Some things, Raz notes, are valuable only instrumentally, as means to other things that are valuable in

¹⁰⁸ Frankfurt, *Necessity*, 152.

¹⁰⁹ Ibid.

¹¹⁰ See Joseph Raz, “On Frankfurt’s Explanation of Respect for People,” in *Contours of Agency*, eds. Sarah Buss and Lee Overton (Cambridge, Mass.: MIT Press, 2002), 301-11.

¹¹¹ Cf. Green, “Two Worries,” 213.

¹¹² Joseph Raz, *Value, Respect, and Attachment* (Cambridge: Cambridge University Press, 2001), 144.

themselves. But even those things that we take to be valuable in themselves (say, works of art) are still only conditionally valuable, since they are valuable *for* whatever can value them—namely, people.¹¹³ Raz claims that not everything can be only conditionally valuable; at least something in the world must have unconditional value if anything else is to have any value at all. For if all valuable things were only conditionally valuable, then there would be nothing for them to be valuable for, nothing to appreciate and engage with their value.¹¹⁴ According to Raz, it is persons who are of unconditional value and whose existence thus forms the “ultimate link in the chain” of value, endowing other objects and beings with (conditional) value.¹¹⁵ More specifically, because they are “valuers,” intentional beings who can act for the “(perceived) normative aspects of the world” embodied in reasons, persons provide the *raison d’etre* for everything else of value.¹¹⁶ This, Raz contends, is how we can best understand Kant’s observation that people are “ends in themselves.”

Raz believes that persons are entitled to respect in virtue of their status as valuers, beings of ultimate value. On Raz’s account, anything valuable is entitled to respect in virtue of whatever gives it value; respecting things of value is one important way in which we appropriately respond to value. More specifically, Raz identifies three “stages of correct response to value”: first, “appropriate psychological acknowledgment of value, that is, regarding objects in ways consistent with their value, in one’s thoughts”; second, preserving what is of value and refraining from destroying it; and third, “engag[ing] with value in appropriate ways.”¹¹⁷ Duties to respect value concern only the first and second of these stages, requiring us to acknowledge value and to preserve it and refrain from

¹¹³ Ibid., 145-47.

¹¹⁴ Ibid., 147, 150-51.

¹¹⁵ Ibid., 157.

¹¹⁶ Ibid., 152, 156.

¹¹⁷ Ibid., 161-62.

destroying it, but not to engage with it by participating in activities that realize it.¹¹⁸ So, too, with respect for persons: we respect persons by acknowledging their being of ultimate value, helping to preserve that value, and refraining from destroying it. On Raz's account, then, respect for persons differs from respect for value *tout court* not in terms of the general kinds of attitudes and actions it requires, but in terms of the unique kind of value that humans possess (*viz.*, ultimate value) and the consequently unique and especially stringent duties that that value grounds.¹¹⁹

This account of respect for persons implies that duties of respect include duties to have certain attitudes—namely, to regard persons as possessing ultimate value. Having such attitudes is simply the first stage of responding to value. And yet, Raz insists that “respecting people is a way of treating them. It is neither a feeling, nor an emotion, nor a belief It is a way of conducting oneself, and more indirectly, of being disposed to conduct oneself towards the object of the respect.”¹²⁰ How can we reconcile this statement with Raz's account of the first stage of responding to value, which seems to involve certain attitudes?¹²¹ Raz must mean not that respect for persons does not consist in attitudes *at all*, but that respect for persons does not consist in attitudes *alone*; even if having appropriate attitudes is not sufficient for respecting persons, having such attitudes must, on Raz's account, be necessary. Indeed, even in the quoted statement, Raz claims that respect for persons partly consists in “*being disposed* to conduct oneself towards the object of the respect,” and in later work, he describes respect for persons as a “stance” amounting to “openness toward value,” which includes “the attitudes that are apt toward

¹¹⁸ Ibid., 164.

¹¹⁹ Ibid., 169-70.

¹²⁰ Ibid., 138.

¹²¹ This tension is also noted by Leslie Green in “Two Worries,” 222.

value.”¹²² Although the ideas of “being disposed” and displaying “openness” are admittedly ambiguous, Raz makes clear that they refer to attitudes as well as treatment.

If one must have certain attitudes in order fully to respect another person as a person, then one can disrespect that person by harboring contrary attitudes, attitudes that regard that person as not possessing ultimate value. But what if one simply lacks any attitudes regarding a person? Does one thereby disrespect him? Raz’s account suggests not. For Raz seems to believe that one need not have respectful attitudes toward something of value if it has yet to enter one’s thoughts and one has no reason to think about it.¹²³ Having thoughts about something of value, in other words, “triggers” duties to have *appropriate* thoughts about it, to respect it.¹²⁴ This position, however, conflicts with our intuition that one can disrespect another person not just by actively scorning him, but also by passively disregarding him. As Leslie Green eloquently puts it, “[t]houghtlessness is one of the familiar modalities of disrespect in the modern world.”¹²⁵ To avoid disrespecting other persons, then, one must cultivate respectful attitudes as well as purge disrespectful ones.

Although Raz recognizes that we have moral reasons to cultivate at least some attitudes toward others, he denies that those reasons amount to genuine duties, and he equivocates about whether they are reasons of respect (as opposed to some more generic moral principle).¹²⁶ But even if we have no duties to cultivate respectful attitudes toward a person, it seems that we do have duties to refrain from treating him in ways that logically entail a *lack* of such attitudes, that logically commit us *not* to regarding him as possessing ultimate value (whatever else we might think, or however else we might feel,

¹²² Joseph Raz, “On Respect, Authority, and Neutrality: A Response,” *Ethics* 120 (2010): 284.

¹²³ See Raz, *Value*, 161 (“There is no general reason to know, or believe, that what is of value is of value. . . . But there is a general reason that if we think of an object which is of value, we should think of it in ways consistent with its value.”).

¹²⁴ Green, “Two Worries,” 220.

¹²⁵ *Ibid.*, 222.

¹²⁶ Raz, “On Respect,” 282 n.4.

about him). Because the relevant attitudes are attitudes of respect, such duties are duties of respect.

Discrimination as Disrespect

According to the account of discrimination embodied in antidiscrimination law, one discriminates against another person based on race or sex when one takes her race or sex as a reason to treat her less favorably than one would treat people of other races or the other sex. We are now in a position to see why acting on this intention is not only illegal, but immoral. Put simply, acting on a discriminatory intention is disrespectful. More specifically, taking another person's race or sex as a reason to treat her less favorably entails lacking a respectful attitude toward her, failing to regard her as possessing ultimate value.

As we have seen, persons are of ultimate value because they are valuers, intentional beings who can act for the reasons given (directly or indirectly) by values.¹²⁷ This capacity endows persons with a great "potential": "the potential to engage with value in the right way, and be thereby enriched or improved, etc."¹²⁸ This potential, in turn, imposes on us a fundamental duty: always to recognize other persons in our rational deliberations as beings whose flourishing matters, whose lives stand to go well or badly depending on their engagement with value.

Discrimination precludes such an attitude. When one discriminates based on race or sex, one takes the victim's race or sex as a reason to exclude her—but not members of other races or the other sex—from some activity or good that one regards as an opportunity for engaging with value or facilitating its realization. Acting on this intention commits one to believing (even if one does not consciously believe) that the flourishing of the victim and the other members of her race or sex matters less than the flourishing of

¹²⁷ Raz, *Value*, 153.

¹²⁸ *Ibid.*, 157.

members of other races or the other sex, that it is more important for the latter to engage with the relevant value than it is for the former to do so. To be sure, many facts about a person will preclude his engagement with certain values. Some people simply lack the cognitive capacities or technical skills necessary to engage with the values realized in, say, reading a novel or building a house. But despite this variation among people's abilities and other traits, we suppose that everyone's life stands to be enriched through engagement with every value and, consequently, that one person's engagement with value is no more important than any other's. One who discriminates based on race or sex, by contrast, is committed to believing that the victim and the other members of her race or sex lack the same relationship to value that people who do not share her race or sex enjoy. This belief is incompatible with an attitude that regards *every* person as having ultimate value—that is, as being a valuer, the kind of being for whom all values present opportunities for flourishing. Because respect for persons partially consists in having such an attitude, the discriminator's evident lack of that attitude constitutes disrespect.

But is it really plausible to think that *every* act of discrimination disrespects its victim, that discriminatory intentions *necessarily* preclude properly respectful attitudes? Against this claim, it might be argued that we cannot assess the (im)morality of a discriminator's intentions without also considering his motives, the further reasons that prompt him to take another person's race or sex as a reason to treat her less favorably. As noted above in Section IV, the antidiscrimination regimes of contemporary liberal democracies prescind from discriminators' motives, condemning equally everyone who acts on discriminatory intentions, irrespective of his precise rationale for doing so. And yet, we might criticize this approach as blunderbuss, as insufficiently attentive to different discriminators' precise circumstances. In particular, we might think that a person can

discriminate intentionally yet out of necessity—and even regretfully—and that such a person does not merit the epithet “disrespectful.”

To test this suggestion, consider the following example. John owns a barbershop in 1960s Jim Crow Mississippi. Immune to his corrupted social milieu, he befriends Bill, a black man, whom he decides to employ as a barber. John’s exclusively white clients, however, are incorrigible racists and have threatened to cease patronizing his barbershop unless he fires Bill. If they carried through on their threat, John would have to close his barbershop and would consequently suffer severe economic deprivation. With great pain and regret, and expressing a profound sense of shame, John fires Bill.

Does John thereby disrespect Bill? Contemporary antidiscrimination law would clearly deem John’s action impermissible discrimination, since John takes Bill’s race as a reason to treat Bill less favorably than he would treat one of his other, white employees. From the perspective of contemporary antidiscrimination law, it matters not that John discriminates against Bill for the undeniably weighty purpose of saving his business and thus his economic livelihood. And yet, we seem loath to pronounce John’s conduct “disrespectful”: while he discriminates against Bill “because of” Bill’s race, he harbors no animus toward Bill or black people more generally; on the contrary, he experiences severe regret and shame at his conduct and expresses that regret and shame to Bill. Though perhaps lamentable, such conduct seems hardly to constitute disrespect. The example thus suggests that discrimination is at least sometimes perfectly permissible and that it need not disrespect its victims.

The example, however, warrants neither conclusion. At most, the example shows that discrimination is not always wrongful *all things considered* and that otherwise disrespectful conduct can sometimes be justified. Discrimination, in other words, might be a pro tanto wrong and a pro tanto form of disrespect. In the example, even if we think

that John was justified in discriminating against Bill in order to save his business, his conduct is still problematic. For John has sacrificed another person's well-being—another person's engagement with value—for his own simply because of the other person's race. Such a decision is disrespectful because it precludes an attitude that regards Bill as bearing the same relationship to value as white people. By capitulating to his racist clients, John at least implicitly buys into their racist logic, privileging his own economic well-being over Bill's simply because Bill is black. To be sure, John does not share his clients' racial animus, and he even experiences regret and shame at his decision. He thus is no racist. But one can disrespect people of other races without being a racist, without consciously believing that people of other races stand to gain less from engagement with value. So long as one makes a decision that *commits* one to this belief, one's decision constitutes disrespect. Thus, to the extent that John's decision has this implication, it is disrespectful and thus wrongful, even if the disrespectful and wrongful features of his decision are overridden by the necessities of his economic circumstances.

This conclusion becomes even more plausible once we consider how we expect John to react emotionally to his decision. Even if we think John justifiably fires Bill, we think he rightly feels pain, regret, and shame at having to do so. He should not rejoice in, or even be nonchalant about, his decision. But why are John's emotional reactions rationally justified rather than pathological? We might think John rightly reacts as he does only because Bill is his friend. But even if Bill were not his friend, we would still expect John to experience some regret and shame at his decision (though perhaps to a lesser degree) and to express this shame and regret to Bill. We would expect John to recognize, in other words, that Bill is not being treated as he should be. This intuition suggests that John's decision contains some wrong-making feature that, while perhaps overridden by other considerations in some circumstances, nevertheless taints John's decision at least to

some degree. Otherwise, John would have no reason to experience the emotions we expect him to.

It is important to note that, on the account developed here, acting on discriminatory intentions is wrong not because such intentions evidence the actual presence of disrespectful attitudes, but because they entail the absence of respectful ones. Thus, *pace* popular accounts of the wrongfulness of discrimination,¹²⁹ it matters not whether the discriminator actually harbors a “prejudiced” belief about his victim, a belief that the victim occupies a lower moral status than people of other races or the other sex. All that matters is that the discriminator acts on an intention that precludes an attitude of respect for the victim. Whether or not we have duties to cultivate respectful attitudes, we certainly have duties not to act on intentions that entail the lack of such attitudes, that are inconsistent with regarding the persons toward whom we act as beings of ultimate value.

Once we appreciate the distinction between actually having disrespectful attitudes, on the one hand, and acting on intentions that entail the absence of respectful ones, on the other, we can answer some of the objections that have been leveled against respect-based accounts of the wrongfulness of discrimination. For example, Kasper Lippert-Rasmussen argues that such accounts appear to condone forms of discrimination that we intuitively condemn—in particular, forms of discrimination where the discriminator acts from a brute desire rather than a belief or value judgment about the victim’s moral worth or where the discriminator discriminates unconsciously, perhaps even acting against a conscious belief or value judgment that the victim possesses as much moral worth as people of other races or the other sex.¹³⁰ But as we have seen, antidiscrimination law prescind from the further desires or beliefs with which people act; one need only take

¹²⁹ See, e.g., John Hart Ely, *Democracy and Distrust* (Cambridge, Mass.: Harvard University Press, 1980), 145-70.

¹³⁰ See Kasper Lippert-Rasmussen, “The Badness of Discrimination,” *Ethical Theory and Moral Practice* 9 (2006): 179.

another person's race or sex as a reason to treat her less favorably than one would treat people of other races or the other sex in order to commit unlawful discrimination. Nor need that the intention be conscious, since, again as we have seen, one can exercise a kind of virtual control over one's actions so that they remain intentional. Just like antidiscrimination law, a respect-based account of the wrongfulness of discrimination can condemn discrimination where the discriminator acts without a prejudiced belief or value judgment—and even where the discriminator acts against a conscious, avowed belief in the moral worth of his victim—so long as the discriminator (consciously or unconsciously) takes race or sex as a reason to treat his victim less favorably. Because such an intention is inconsistent with a properly respectful attitude, acting on it is disrespectful and therefore immoral.

More significantly, Lippert-Rasmussen argues that respect-based accounts of the wrongfulness of discrimination have a counterintuitive implication: that it is morally worse to discriminate with a false belief about the moral worth of one's victim than to discriminate with a true belief.¹³¹ If a respect-based account locates the wrongfulness of discrimination in the beliefs or attitudes of discriminators, then it seems that it must condemn, say, one who discriminates against a black person while falsely believing that black people are morally inferior to white people more severely than it would condemn one who discriminates against a black person while rightly believing that black people and white people are morally equal. And yet, we tend to think that the second discriminator acts more immorally than the first, that the second quite literally “knows better.”

Once again, however, Lippert-Rasmussen's objection misidentifies the feature of discriminatory acts that respect-based accounts (or at least the one propounded in this

¹³¹ Ibid., 183-84.

chapter) condemn. Whereas Lippert-Rasmussen depicts respect-based accounts as condemning the beliefs with which discriminators act, the account advanced in this chapter condemns the logical implications of acting on discriminatory intentions—namely, the incompatibility of such intentions with an attitude that regards every person as possessing ultimate value. This alternative focus allows for somewhat complicated cases, such as where a person acts on intentions that entail the lack of a properly respectful attitude yet nevertheless professes to hold such an attitude. There are a number of possible explanations for such a case. Perhaps the person generally does believe that all people are of ultimate value yet nevertheless succumbed to the temptation to discriminate on one particular occasion, revealing his belief to be less than fully secure. Or perhaps his professed belief exists only at the most abstract level, failing to penetrate into his dispositions. Or perhaps he simply fails to grasp the practical implications of his belief. The exact explanation is unimportant. What matters for our purposes is that a respect-based account of the wrongfulness of discrimination focuses on the logical implications of acting on discriminatory intentions rather than people's professed beliefs and thus does not have the counterintuitive implication that Lippert-Rasmussen ascribes to it.

Whereas Lippert-Rasmussen's objections imply that a respect-based account of the wrongfulness of discrimination is underinclusive, we might worry that such an account is actually overinclusive. For we often act on intentions that seem to entail the lack of a properly respectful attitude. When we scorn someone of a lower social status, someone whom we consider unattractive, or someone with different aesthetic tastes, the intention on which we act can often plausibly be thought to preclude an attitude that regards the victim as possessing ultimate value. And yet, the law tolerates such forms of disrespect. What, then, justifies antidiscrimination law's relatively narrow focus on forms of disrespect involving race and sex (in addition to a few other traits)?

There are two responses to the overinclusion concern, each of which invokes history as a kind of limiting principle. The first response concedes that many forms of intentional conduct are disrespectful but notes that, historically, only a few have been sufficiently pervasive to warrant a legal response. We tolerate, in other words, forms of disrespectful conduct that are sufficiently marginal in society in general and the domains of employment and public accommodations in particular either to elude legal detection or to lack a significant claim on the government's limited resources. To be clear, unlike the doctrines of political legitimacy considered in Part I of this thesis, this response condemns the prohibition of much disrespectful conduct because such regulation would be imprudent, not illegitimate.

The second response invokes history in a more fundamental way. Rather than use history to distinguish pervasive from marginal disrespectful conduct, the second response uses history to identify disrespectful conduct in the first place. According to this response, many forms of conduct that we might think are disrespectful in the abstract are not disrespectful in actual social practice. For as Joseph Raz notes, respect has a “symbolic” aspect, with people showing respect or disrespect to one another through various “symbolic acts,” whose meaning, in turn, is defined largely by a society's history and social practice.¹³² (Think, for example, of the significance of doffing one's cap in different societies.) Although Raz invokes history and social practice to identify which conduct we interpret as disrespectful,¹³³ we might think that they also determine which intentions are disrespectful—that is, which intentions, when acted on, preclude a properly respectful attitude. On this view, taking a person's unattractiveness, say, as a reason to treat him less favorably may be callous, unsympathetic, or petty, but it does not become truly disrespectful until history and social practice endow such an intention with a

¹³² Raz, *Value*, 171-72.

¹³³ This issue is considered below.

meaning that represents unattractive people as lacking ultimate value. Put another way, history and social practice might transmute a trait from a basis of contempt toward one particular person who bears it into a basis of disrespect toward all persons who bear it, simply in virtue of their bearing it. This view would, of course, imply that the first instance of discrimination against a black person or woman was not disrespectful. But we should not insist on an ahistorical moralism to explain all immorality, especially in the case of conduct as historically and socially contingent as discrimination based on race or sex.¹³⁴

Even if we condemn discrimination as a form of disrespect, we can still acknowledge that the conduct is often wrong for other reasons as well. As the previous chapters have noted, discrimination frequently (though not always) harms its victims' well-being and insults their self-respect. Thus, as with most disrespectful conduct, the case against discrimination is "rationally overdetermined";¹³⁵ we often have multiple, individually sufficient reasons not to discriminate. But as we also saw in the previous chapters, the harmful effects of discrimination do not always materialize, and in such cases, harm-based reasons do not condemn the conduct. If we wish to condemn every instance of discrimination, then we will have to attend to the feature that makes the conduct intrinsically wrongful, irrespective of its effects—namely, the disrespectful nature of the intentions on which discriminators act.

Disrespectful Intentions, Not Disrespectful Social Meanings

We might worry that the respect-based account of the wrongfulness of discrimination developed above misses the true significance of disrespectful conduct. In particular, whereas that account focuses on the attitudes that acting on certain intentions

¹³⁴ For Raz's analogous view that the intentions of lawmakers are constrained by objective features of linguistic and social context, see his *Between Authority and Interpretation* (Oxford: Oxford University Press, 2009), ch. 11.

¹³⁵ Green, "Two Worries," 226.

logically precludes, we might think that disrespectful conduct is wrong because it tends to hurt its victims. According to Frankfurt, for instance, to disrespect someone is to “assault . . . his [social] reality” and thus to provoke in him “painful feelings of resentment” as well as “a more or less inchoate anxiety.”¹³⁶ Similarly, Raz notes that, unlike non-human objects of value, people “have a sense of their own identity, a sense that they are of value, and therefore are hurt by disrespect.”¹³⁷ For Frankfurt and Raz, honoring our respect-based duties is so urgent not because disrespectful conduct reflects badly on our moral characters, but because such conduct profoundly insults its victims.

To focus on the hurt caused by disrespectful conduct, however, is to identify respect too closely with people’s subjective emotional reactions. In particular, when applied to the issue of discrimination, such a subjective account seems to encounter precisely the same impediment that thwarted the harm principle—namely, an inability to condemn discriminatory acts that happen not to have deleterious consequences for their victims. Raz himself appears to appreciate this concern, criticizing Frankfurt’s initial formulation of duties of respect as too subjective because it “shows what hurt is caused to people when they *think* that they are treated wrongly, not when they actually are treated wrongly.”¹³⁸ By defining disrespectful treatment in terms of its subjective emotional effects rather than its more objective features, Frankfurt reduces it to a mere epiphenomenon, mistaking people’s reactions to a moral wrong for the moral wrong itself. Such a subjective account threatens to conflate our ostensibly distinctive duties of respect with our more generic moral obligations, which already enjoin us to avoid hurting others—whether through disrespectful conduct or otherwise.¹³⁹

¹³⁶ Frankfurt, *Necessity*, 152-53.

¹³⁷ Raz, *Value*, 171.

¹³⁸ Raz, “Frankfurt’s Explanation,” 308.

¹³⁹ Raz, *Value*, 171-72.

We might begin to abstract from the subjective perspective of the victims of disrespectful treatment by considering a more objective one: that of a reasonable social interpreter. For example, Raz argues that certain actions and social practices function as “symbolic expressions of respect or disrespect,” conveying meanings that “reasonable” interpreters would understand as respectful or disrespectful.¹⁴⁰ According to Raz, so long as an action or practice carries an objectively disrespectful meaning in a particular historical and cultural context, so long as an objective member of the society would understand the action as denying the value of its victim, it is disrespectful.

It is from the perspective of the objective interpreter that Deborah Hellman condemns discrimination based on race or sex. According to Hellman, discrimination is wrong insofar as it “demeans” its victims, conveying a social meaning that implies that they possess less moral worth than other persons.¹⁴¹ Because social meaning, like language, is largely “conventional,”¹⁴² a form of treatment will convey such a meaning only if various features of the surrounding social “context” and “culture” render it such.¹⁴³ Therefore, just as we must make an “interpretive judgment” to distinguish orders from advice, so, too, must we construe context and culture to distinguish demeaning treatment from less egregious or even perfectly innocuous behavior.¹⁴⁴

At least when perpetrated against racial minorities and women, Hellman argues, discrimination is demeaning. Like any other kind of conduct, discrimination based on a particular trait can demean its victims only if historical practice and current social context imbue it with a disrespectful connotation. Discrimination will tend to garner such a connotation if it targets social groups whose members have historically endured

¹⁴⁰ Ibid.

¹⁴¹ Hellman, *When Is Discrimination Wrong?*, 8, 29, 48. For a similar account, see Patrick S. Shin, “The Substantive Principle of Equal Treatment,” *Legal Theory* 15 (2009): 149-72.

¹⁴² Hellman, *When Is Discrimination Wrong?*, 35.

¹⁴³ Ibid., 7-8, 38.

¹⁴⁴ Ibid., 61-62.

systematic mistreatment and who now occupy a disadvantaged social status.¹⁴⁵ Thus, given its legacy of racial oppression and the present deprivation of much of its black population, the United States today provides a social milieu in which discrimination against black people will almost always demean its victims.¹⁴⁶ So, too, will most instances of discrimination against women, who have been historically subjugated and are currently disadvantaged.¹⁴⁷ Possessing no comparable historical or contemporary significance, discrimination based on other arbitrary traits like eye color will rarely demean its victims (though it might violate other moral principles).¹⁴⁸ Hellman thus distinguishes discrimination against racial minorities and women from other, relatively unobjectionable forms of conduct based on the former's tendency to demean its victims.

But if we follow Hellman in focusing on social meaning, we will have to revise some of our moral intuitions about discrimination based on race or sex. For many instances of the conduct imply no demeaning connotation. The kind of covert discrimination considered in the previous section, for example, will rarely demean its victims, since its perpetrators tend to devise plausible subterfuges to dissemble their discriminatory purposes. Consider once again the employer who uses state residency as a proxy for race. Assuming that state residency also correlates with a legitimate characteristic like educational attainment and that the employer's disavowals of discriminatory intentions are credible, an objective interpreter will understand his discriminatory actions not as a demeaning affront to his victims' dignity, but as a permissible business practice.¹⁴⁹ And yet, we tend to deem his conduct seriously immoral even though it conveys no demeaning connotation. To be sure, were other people to discover and expose the employer's true aims, his actions would then indeed acquire such

¹⁴⁵ Ibid., 24-25.

¹⁴⁶ Ibid., 28, 40.

¹⁴⁷ Ibid., 15.

¹⁴⁸ Ibid., 14.

¹⁴⁹ Hellman seems to concede this implication of her theory. See *ibid.*, 160.

a meaning. But why should we attempt to uncover discriminatory intentions in the first place? If, like Hellman, we located the immorality of discrimination in its social meaning, then we would have no reason to expose covert instances of discrimination, whose disrespectful connotation remains occluded to an objective interpreter (who, as we have seen, lacks access to the perpetrator's intentions). On the contrary, it seems, somewhat perversely, that we would have reason to *conceal* the discrimination, lest we actuate what, on Hellman's account, remains a merely latent wrong.

As we saw in the previous section, Hellman is led to this counterintuitive conclusion by her refusal to consider the intentions of discriminators, the reasons for which they act. In this respect, Hellman's account goes even further than Raz's, which apparently declines to divorce social meaning from people's intentions (though he believes that intentions are themselves largely constrained by social context).¹⁵⁰ As Raz seems to appreciate, discrimination is not wrong because it tends to convey a disrespectful connotation; rather, it tends to convey a disrespectful connotation because it is wrong. It is wrong, moreover, because of the intentions on which discriminators act. Hellman, by contrast, mistakes a mere epiphenomenon for the underlying phenomenon itself. Her theory thus grounds a duty to perform what Raz dubs "symbolic expressions of respect" but no duty to be genuinely respectful. If we want to ground this latter, more fundamental duty, if we want to oblige people not merely to be seen to show respect but actually to practice it, then we must consider people's intentions, the reasons for which they act.

Considering Hellman's account of the wrongfulness of discrimination also helps us to distinguish the respect-based account developed in this chapter from Denise Réaume's dignitarian account, considered in Chapter 3. First, and least significant,

¹⁵⁰ See Green, "Two Worries," 229.

Réaume grounds human dignity in a different aspect of personhood: whereas this chapter's respect-based account follows Raz in grounding duties of respect in persons' "ultimate value," Réaume grounds their dignity in their "autonomy," their capacity to make choices and to have a sense of self.¹⁵¹ This difference, however, seems to be more theoretical than practical, for each aspect of personhood can be (and has been) thought to entail duties of respect.

Second, and more important, Réaume's account attends to victims' subjective reactions to discrimination, whereas this chapter's account prescind from them. To be sure, Réaume purports to propound a purely objective account of the wrongfulness of discrimination, one that condemns the conduct regardless of the effects it has on its victims. But as Chapter 3 demonstrated, Réaume compromises the objectivity of her account by repeatedly adverting to subjective notions such as "self-esteem" and "humiliation." Although this chapter's account by no means discounts the significance of the subjective psychological effects discrimination can have for its victims, it does attempt to explain why discrimination is wrongful even when such effects do not materialize.

Third, even if Réaume succeeds in propounding a purely objective account of the wrongfulness of discrimination, her account still differs from the one developed in this chapter, which is self-consciously subjective in another respect. In particular, it attends to certain subjective features of the perpetrators of discrimination—namely, their intentions and attitudes. Réaume, by contrast, claims that discrimination is wrongful not in virtue of the intentions of its perpetrators, but solely in virtue of its objective social meaning, which is completely independent of discriminators' intentions.¹⁵² Réaume, in other words, believes that discrimination can denigrate its victims even when not perpetrated

¹⁵¹ Denise G. Réaume, "Discrimination and Dignity," *Louisiana Law Review* 63 (2003): 673-77.

¹⁵² Denise G. Réaume, "Harm and Fault in Discrimination Law: The Transition from Intentional to Adverse Effect Discrimination," *Theoretical Inquiries in Law* 2 (2001): 365-66.

intentionally. In this respect, Réaume's account resembles Hellman's more than the one developed in this chapter and thus suffers from similar shortcomings. Moreover, as we shall see in the next chapter, it is precisely its focus on discriminators' intentions that renders this chapter's account perfectionist, perfectionism that, as we saw in Chapter 3, Réaume eschews.

These distinctions should not, of course, obscure the affinities this chapter's account shares with Réaume's. Both, for instance, are motivated by a desire to condemn all instances of discrimination, even those whose perpetrators are neither racist nor sexist, but who instead discriminate for more prosaic reasons, such as economic efficiency. And both ground antidiscrimination principles in intrinsic features of our personhood rather than more contingent features of contemporary society. But ultimately, despite their shared use of terms such as "dignity" and "respect," the two accounts remain irreconcilable, particularly because of this chapter's account's focus on discriminators' intentions and its attendant perfectionism.

Implications

This subsection offers some tentative thoughts about three potential implications of the respect-based, intention-focused account of the wrongfulness of discrimination developed in this chapter. First, the antidiscrimination laws of contemporary western liberal democracies are symmetric, forbidding discrimination against men and members of the racial majority just as readily as they forbid discrimination against women and members of racial minorities. Whereas a harm-based account of discrimination understands this feature as a mere political compromise, a respect-based account understands it as a deeply principled reflection of morality. For according to the latter account, it is wrong to take a person's *race* or *sex*, whatever it may be, as a reason to treat her less favorably than one would treat people of other races or the other sex. Even when

the victim is white or a man, such an intention entails the lack of an attitude that regards the victim as possessing ultimate value and is therefore disrespectful. A respect-based account thus disagrees with many contemporary liberal philosophers, who tend to believe that discrimination is wrong only when perpetrated against members of historically disadvantaged groups.¹⁵³ Such a belief is belied not only by antidiscrimination law itself, but also by the fact that members of disadvantaged groups can regard (or be committed to regarding) members of advantaged social groups or members of their own groups as lacking ultimate value, as being something less than full persons.¹⁵⁴

Second, while a respect-based account confirms the symmetry of antidiscrimination law, it appears to condemn the provision of U.S. antidiscrimination law that permits employers to discriminate based on sex whenever sex is a “bona fide occupational qualification” (“BFOQ”).¹⁵⁵ Suppose, for example, that a theater company takes a male actor’s sex as a reason not to cast him in a female role. It seems that the theater company commits discrimination according to the definition enshrined in contemporary antidiscrimination law and that its conduct is disrespectful according to a respect-based account of the wrongfulness of discrimination. To accommodate our intuitions about this case, it appears that we must follow antidiscrimination law in carving out an ad hoc exception from the general definition of discrimination. Although the precise scope of this exception is unclear (in law as well as morality), it undoubtedly should encompass aesthetic performances. Admittedly, this kind of ad hoc approach is unsatisfying as a matter of principle, but it seems necessary to address the unique ethical dilemmas that aesthetics presents.

¹⁵³ See, e.g., Scanlon, *Moral Dimensions*, 73-74.

¹⁵⁴ Cf. Joshua Glasgow, “Racism as Disrespect,” *Ethics* 120 (2009): 89 (arguing that members of powerless racial groups can still be racist toward members of powerful ones).

¹⁵⁵ Title VII of the Civil Rights Act of 1964, § 703(e)(1). Interestingly, according to this provision, race can never be a BFOQ.

Also problematic are cases where sex is arguably a BFOQ only because it is a “reaction qualification.”¹⁵⁶ Suppose, for instance, that women generally prefer to be examined by female gynecologists due to privacy concerns. May a medical practice therefore refuse to hire male gynecologists? U.S. courts are divided on this issue, as, it seems, are our intuitions. On the one hand, many women may very well have legitimate privacy concerns about being examined by male doctors. On the other hand, it is unclear why the content of customers’ (or patients’) preferences should determine the legality and morality of employers’ discriminating based on sex. Indeed, taken to its logical extreme, a theory that caters to people’s privacy-based preferences would hold that no men could become gynecologists if all women wanted to be examined by other women (or that no women could become urologists if all men wanted to be examined by other men). Perhaps we should nevertheless adopt another ad hoc exception to cover this case. But it is by no means obvious that a respect-based account should be embarrassed for condemning as disrespectful employers’ discriminatory efforts to cater to their customers’ privacy-based preferences.

Third, although this thesis cannot explore the issue in depth, it appears that a respect-based account of the wrongfulness of discrimination condemns “affirmative action” (or “positive discrimination”) programs—whether the justification for such programs is compensating for past group-based injustice,¹⁵⁷ promoting diversity,¹⁵⁸ or overcoming segregation.¹⁵⁹ For in order to take some people’s race or sex as a reason to treat them more favorably, we must take other people’s race or sex as a reason to treat

¹⁵⁶ Wertheimer, “Jobs.”

¹⁵⁷ See, e.g., Judith Jarvis Thomson, “Preferential Hiring,” *Philosophy & Public Affairs* 2 (1973): 364–84. By contrast, a respect-based account would not condemn programs designed to compensate the actual victims of discrimination, since such compensation would be based on their prior victimization, not their race or sex.

¹⁵⁸ See, e.g., Ronald Dworkin, *Sovereign Virtue* (Cambridge, Mass.: Harvard University Press, 2000), 386–408.

¹⁵⁹ See, e.g., Elizabeth Anderson, *The Imperative of Integration* (Princeton: Princeton University Press, 2010).

them less favorably; ordinary discrimination is a necessary means for positive discrimination. A university, for instance, cannot take an applicant's being black as a reason to admit him and then act on that intention without taking at least one other equally well-qualified applicant's being white as a reason to reject him and then acting on that intention (assuming a scarce number of positions and typical demographics). We might nevertheless continue to condone affirmative action, treating discrimination as a pro tanto wrong that can be outweighed by appeals to compensatory justice, diversity, or integration. Whether such appeals actually outweigh the wrong of discrimination depends on which general normative theory we happen to espouse. But at least within the domain of respect, it is always wrong to take a person's race or sex as a reason to treat her less favorably than one would treat people of other races or the other sex, irrespective of one's ulterior purposes.

VIII. CONCLUSION

To justify the uniform, continuous antidiscrimination regimes of contemporary western liberal democracies, we must construe antidiscrimination law as a measure not just for providing opportunities to discrimination's victims, but also for regulating the intentions of discriminators, the reasons for which they act. Antidiscrimination law itself adopts precisely this cognitive focus, defining discrimination in terms of people's intentions. In so doing, antidiscrimination law merely tracks morality, which also predicates the permissibility of many of our actions on the reasons for which we act.

Discriminatory intentions are so objectionable because they preclude a properly respectful attitude toward the victims. Neither a subjective insult inflicted on victims nor an objective meaning perceived by observers, disrespect is a product of the reasoning of discriminators. More specifically, by taking another person's race or sex as a reason to treat her less favorably than one would treat people of other races or the other sex, one

necessarily commits oneself to regarding her as lacking the ultimate value we suppose all persons to possess. Only if we eschew discrimination, treating black people as favorably as we would white people and women as favorably as we would men, can we accord every person the respect she deserves.

We cannot legitimize antidiscrimination law, however, simply by condemning discrimination as a form of disrespect. Rather, we must first construct a doctrine of political legitimacy that authorizes the government to regulate at least some disrespectful conduct. Unless it may at least occasionally prevent its citizens from disrespecting one another, the government will have to tolerate many violations of respect, including many instances of discrimination. But how intrusive must governmental regulation be in order to prevent violations of respect? Does a respect-based doctrine of political legitimacy still accord with the fundamental tenets of liberalism? Or does it presuppose a thoroughgoing perfectionism, licensing the government to cultivate the moral characters of its citizens? The final chapter considers these questions.

Chapter 6

Respect and Politics

I. RESPECT, INTENTION, AND THE POWER OF THE LAW

Despite its significant shortcomings, contemporary liberalism teaches an important lesson about the proper limits of political power, a lesson that even political perfectionists should heed. In particular, contemporary liberals demand that we always distinguish the question of political legitimacy from the question of moral permissibility, the question of whether a particular form of conduct may be legally proscribed from the question of whether it should be morally deprecated. To conflate these questions and to infer the propriety of prohibiting a particular form of conduct directly from its immorality, contemporary liberals rightly worry, is to neglect other important considerations relevant to any decision to exercise political power. Contemporary liberalism thus errs not necessarily in attempting to delineate a set of values that are especially salient in the political realm, but rather in purporting to sever such values from more generic moral principles—a blunder whose considerable costs become all too clear when we consider liberalism's position on antidiscrimination law. If we are to legitimate the full extent of antidiscrimination law, then we must deny any strict dichotomy between politics and the rest of morality, grounding our justification of the legislation in the fundamentally immoral feature of discrimination based on race or sex identified in the previous chapter—namely, the lack of respect entailed by the intentions of the conduct's perpetrators. But in construing antidiscrimination law as a measure for enforcing respect, we must also be careful not to disregard the discontinuities between politics and morality identified by contemporary liberalism, lest we allow our antidiscriminatory zeal to overwhelm our justified suspicion of political power. Legal moralism without moral

fundamentalism—this is the balance we must strike in enlisting antidiscrimination law to enforce respect.

Foremost among the many discontinuities between politics and the rest of morality is the former's apparent focus on people's external actions rather than their internal motivations, a focus that initially seems to preclude any intention-based account of antidiscrimination law. Whereas morality scrutinizes people's intentions to evaluate the actions they perform, the law appears largely to prescind from the reasons for which people act, focusing instead on the external effects of their conduct and their awareness of those effects. Nor should this focus of the law surprise us. For it seems that the law simply cannot require more than merely external conformity with its injunctions, more than the performance or forbearance of certain physical acts.¹ Were the law to go further and enjoin people not merely to comport themselves in a particular manner, but also to act for certain reasons, it would apparently be issuing an idle injunction. Though it may address itself to its subjects' reason, the law can evidently constrain only their behavior.

This is not to suggest that the law necessarily prescind from all our mental states. On the contrary, the law can consider our intentions in the loose sense in which many tort scholars use that term—that is, our awareness of the acts we perform and the effects those acts have for others. Consider, for instance, the law against murder, which defines the offense of murder not just causally, in terms of the changes an agent effects in the external world, but also, through phrases like “malice aforethought,” mentally, in terms of the agent's awareness of those effects. At the very least, then, the law can and does require its subjects to forgo performing certain acts with (actual or imputed) awareness of those acts' effects. But to forbid performing certain actions with certain mental states is

¹ This kind of practical argument is often advanced by liberal theorists, particularly of a Kantian persuasion, as a decisive objection to perfectionism. See, e.g., Christine M. Korsgaard, *Creating the Kingdom of Ends* (Cambridge: Cambridge University Press, 1996), 20; Arthur Ripstein, *Force and Freedom* (Cambridge, Mass.: Harvard University Press, 2009), 269; and T.M. Scanlon, *Moral Dimensions* (Cambridge, Mass.: Harvard University Press, 2008), 59-60, 61 n.15.

not to require people to act for certain reasons. Despite its occasionally assessing people's mental states, the law generally focuses on their external conduct, regulating what they do rather than what they think.

If the law in general cannot compel people to act for certain reasons, then how can antidiscrimination law in particular compel people to be respectful? Recall that the account of discrimination expounded in the previous chapter condemned the conduct for the lack of a properly respectful attitude entailed by the race- and sex-based intentions on which discriminators act. Such an account of discrimination suggests a corresponding account of antidiscrimination law according to which the legislation seeks not only to ameliorate the material privation and other disadvantage of the victims, but also to induce the perpetrators to treat their fellow human beings with respect. But the law seems incapable of executing this latter mission. For the law can instill respect in its subjects only if it can shape their rational deliberations, a degree of internal influence that, as we have seen, the law cannot exert. A perfectionist, respect-based account of antidiscrimination law thus seems to fail on its own terms.

Indeed, given the law's external focus and its seeming indifference to people's reasons for acting, it appears we should reconsider not only perfectionist accounts of antidiscrimination law in particular, but political perfectionism more generally. For that ideology looks to the law to mold people's moral characters by influencing their intentions. If the law cannot exert such influence, if it must content itself with merely constraining people's external actions, then political perfectionism loses its *raison d'être*, largely ceding the ideological battlefield to other, non-perfectionist theories that in no way presume to influence people's rational deliberations. Only if the law reaches people's reason can political perfectionism present a coherent alternative to contemporary liberalism.

Is political perfectionism—and, with it, any perfectionist account of antidiscrimination law—really so misconceived? Is its faith in the capacity of the law to influence people’s rational deliberations and shape their intentions and characters really so misplaced? This chapter will attempt to allay these doubts. First, this chapter will show that the respect-based account of the wrongfulness of discrimination developed in the previous chapter does indeed require a genuinely perfectionist account of antidiscrimination law. In particular, Section II will identify two connections between disrespectful intentions and moral character, one indirect and the other direct. The first, indirect connection is that one blemishes one’s moral character by acting on disrespectful intentions. Thus, while antidiscrimination law, like any other legal measure, cannot ensure that its subjects actually evince respect, that they actually regard their fellow human beings as possessing ultimate value, it can prevent them from performing certain actions while lacking a properly respectful attitude and thus from (further) compromising their characters.

This indirect connection between disrespectful intentions and moral character, however, does not entail perfectionism, for while antidiscrimination law might have the *effect* of preventing people from compromising their characters, it need not adopt this effect as its *purpose*. By contrast, the second, direct connection between disrespectful intentions and moral character does entail perfectionism. To recognize the wrongfulness of discrimination, as defined in the previous chapter, antidiscrimination law must not merely identify the reasons for which people act, but also assess the moral significance of acting on those reasons, as well as the attitudes (or lack thereof) they entail. What attitudes one possesses or lacks, how one does or does not regard others—these features partly constitute one’s character. Antidiscrimination law must attend to these features in

order to proscribe all instances of discrimination, and it can attend to these features only by espousing perfectionism.

But can we not avoid this conclusion simply by declaring discrimination to be an *injustice*, a violation of its victims' *rights*? As Section III will suggest, even if discrimination violates its victims' rights and is therefore unjust, we cannot define the right to respect that discrimination violates without defining that right's correlative duties. And given the nature of respect—in particular, its dependence on people's attitudes—we cannot define duties of respect without considering the duty bearers' moral characters. Even if we need not espouse perfectionism to define many other duties, we must do so to define duties of respect, the kinds of duties that antidiscrimination law enforces.

Sections IV and V will consider some of the more political implications of the version of perfectionism outlined in the previous sections. Can a perfectionist, respect-based doctrine of political legitimacy recognize prudential limits to the law, or is it committed to a kind of moral fundamentalism, the idea that “no sin should go unpunished,” whatever the costs? Might such a doctrine recognize even some principled limits to the law? The last two sections will confirm that a modest respect-based perfectionism can accommodate some of the liberal intuitions that motivate our commitment to limited government.

II. RESPECT, INTENTION, AND CHARACTER

Even if we accept the account of the wrong of discrimination outlined in the previous chapter, we still might question whether we must espouse perfectionism in order to authorize the government to proscribe that wrong. Recall that the previous chapter depicted discrimination based on race or sex as a form of disrespect, locating the conduct's wrongfulness in the intentions on which discriminators act—in particular, the attitudes (or lack thereof) entailed by those intentions. Because the wrongfulness of

discrimination is thus defined in terms of discriminators' mental states as well as their external acts, we might think that the law must assess discriminators' moral characters in order to condemn and prohibit their conduct, the kind of assessment that presupposes perfectionism. But as we saw in Chapter 3 and again in Chapter 5, antiperfectionism permits the government to consider many of the mental states with which people act, such as their awareness of their actions and of the effects their actions have on others—that is, their intentions, understood in the thin sense in which the term is used by tort scholars. Indeed, it seems that antiperfectionism can not only consider whether people are aware of their actions, but also identify their intentions, understood in a thicker sense as what they take to be their reasons for action. For example, to observe that one took another person's race or sex as a reason to treat her less favorably than one would treat people of other races or the other sex, the government apparently need not hold one to be at fault. If the government can thus recognize people's discriminatory intentions without having to condemn their characters, then it seems that antiperfectionist liberalism can justify the full extent of antidiscrimination law after all.

To see why this non-moralistic conclusion is in fact unwarranted, we must explore the connections between discriminatory intentions and discriminators' characters. One connection between a person's intentions and his character in no way impugns antiperfectionism. We tend to think of people's characters as being evinced by their actions and reasons. On this view, a person's character comprises his "dispositions to act for certain reasons, that is, to act, and in acting to have certain reasons for acting,"² so that when a person acts for a certain reason, we gain insight into at least one facet of his character. But in fact, a person's actions and reasons not only evince his character, but also constitute it. As John Finnis rightly notes, our intentional choices "last" in our

² Stephen L. Darwall, "Two Kinds of Respect," *Ethics* 88 (1977): 43. Cf. Aristotle, *Nicomachean Ethics*, trans. Christopher Rowe (Oxford: Oxford University Press, 2002), bk. VI chs. 12-13.

characters.³ That is, when one chooses to act for a particular reason, the choice not only demonstrates that one is disposed to make such a choice, but also *makes* one the kind of person who makes such a choice. When, for instance, one takes another person's race or sex as a reason to treat her less favorably than one would treat people of other races or the other sex, one makes oneself a disrespectful person in addition to showing oneself to be such a person.

This connection between disrespectful intentions and moral character suggests a perfectionist function for antidiscrimination law. Insofar as it prevents people from acting on certain disrespectful intentions, antidiscrimination law prevents them from shaping their characters in bad ways, from molding themselves into disrespectful persons. To be sure, because a person can be disrespectful simply in virtue of having disrespectful attitudes or lacking respectful ones, it is immoral even merely to form the kind of disrespectful intention on which discriminators act. The government cannot directly forestall this kind of immorality—even if it were legitimate for the government to try to do so. But the government can prevent a person from acting on the discriminatory intentions he forms, action that would undoubtedly constitute a further wrong—and thus a further blemish on his character—beyond merely having disrespectful attitudes or lacking respectful ones. In preventing people from acting on discriminatory intentions, moreover, the government might indirectly prevent them from forming those intentions in the first place through a kind of Aristotelian “habituation.”⁴ That is, the government might gradually attenuate a person's discriminatory intentions as a motivating force in his deliberations by repeatedly preventing him from acting on them, thereby molding him into a more respectful person, if not a fully respectful one. Whether directly or indirectly,

³ John Finnis, “Intention in Tort Law,” in *Philosophical Foundations of Tort Law*, ed. David G. Owen (Oxford: Oxford University Press, 1995), 243. See also John Gardner, “Wrongs and Faults,” in *Appraising Strict Liability*, ed. A.P. Simester (Oxford: Oxford University Press, 2005), 62.

⁴ See Aristotle, *Nicomachean Ethics*, bk. II ch. 1.

antidiscrimination law can perform the distinctively perfectionist function of shaping people's moral characters

Though distinctively perfectionist, this function does not necessarily render a government that enacts antidiscrimination law perfectionist as well. For antidiscrimination law's tendency to shape people's characters is merely an effect of the legislation. The government need not espouse this effect as its purpose; it need not intend to shape its citizens' characters, even if it foresees that enacting antidiscrimination law will tend to do so. According to most contemporary liberals, the government complies with antiperfectionism so long as it disclaims a perfectionist purpose. What matters for most contemporary liberals, in other words, is not that the government eschew any policy that has the effect of favoring a particular conception of the good life, but that it not intend to promote a conception of the good life (especially through legislation that molds people's characters in conformity with that conception).⁵ Thus, while antidiscrimination law's tendency to shape people's characters might be perfectionist, such a function is not perfectionist in a way that impugns antiperfectionism.

More problematic for antiperfectionism is another connection between discriminatory intentions and discriminators' moral characters, one that does entail a fully perfectionist account of antidiscrimination law. Although we can identify the reasons on which discriminators act without considering their moral characters, we must consider their moral characters in order to assess the moral significance of their acting on those reasons. Indeed, this may be true for all reasons, not just disrespectful ones. For when agents perform an intentional action, an action for a reason, "what the agents believe about the action leads them to do it, and guides their doing of it, all the way to its

⁵ See, e.g., Charles E. Larmore, *Patterns of Moral Complexity* (Cambridge: Cambridge University Press, 1987), 42-47; John Rawls, *Political Liberalism*, paperback ed. (New York: Columbia University Press, 1996), 192-94. But see Alan Patten, "Liberal Neutrality: A Reinterpretation and Defense," *Journal of Political Philosophy* (forthcoming 2011) (arguing that liberal neutrality requires neither neutrality of aim nor neutrality of effect, but rather neutrality of "treatment").

conclusion, and that suggests that they approve of the action, given what they believe about it.”⁶ We thus might think that we cannot properly recognize an action as intentional, as being done for some reason, without attending to the attitudes of the person who performs it. And we might think further that we cannot assess the moral significance of the person’s performing the action for that reason without assessing the moral significance of his attitudes. Because his attitudes partially constitute his character, his general stance toward value and the reasons it generates, we might conclude that we cannot assess the moral significance of his performing the action for a particular reason without assessing his character.

But whether or not we must consider character in order to assess the moral significance of acting on our intentions in general, we certainly must do so in order to assess the moral significance of acting on discriminatory intentions in particular. As we saw in the previous chapter, acting on a discriminatory intention is wrong because such an intention precludes an appropriately respectful attitude toward the victim of the discrimination. Discrimination is wrong, in other words, because one cannot truly regard another person as possessing ultimate value while taking her race or sex as a reason to treat her less favorably than one would treat people of other races or the other sex. This account defines the wrongfulness of the act of discriminating against someone in terms of the wrongfulness of the intentions of discriminators, and it defines the wrongfulness of the intentions of discriminators in terms of the attitudes such intentions preclude.

Whatever other features it might encompass, a person’s character includes his attitudes, his general dispositions to act for certain reasons and to forbear acting for others. An honest person, for instance, is honest not simply because he tells the truth, but because he is the kind of person who is disposed to take the intrinsic moral worth of truth telling,

⁶ Joseph Raz, “On the Guise of the Good,” in *Desire, Practical Reason, and the Good*, ed. Sergio Tenenbaum (Oxford: Oxford University Press, 2010), 116.

rather than some ulterior motive, as his reason for telling the truth. Similarly, a respectful person is respectful not simply because he makes outward signs of respect, but because he is the kind of person who is disposed to take the intrinsic moral worth of showing respect,⁷ rather than some ulterior motive, as his reason for showing respect. We must therefore reference attitudes in order to describe both the honest person and the respectful one.

But whereas we need not reference attitudes in order to describe the characteristic action of the honest person, we must reference attitudes in order to describe at least some of the characteristic actions of the respectful person. We can describe the characteristic action of the honest person—not intentionally concealing the truth—independently of the attitudes that make a person honest: while allowing that a person who refrains from intentionally concealing the truth only, say, to garner a reputation for truthfulness is not an honest person (i.e., does not display the virtue of honesty), we can nevertheless recognize his action as truthful.⁸ The characteristic actions of the respectful person, by contrast, are partly defined by respectful attitudes. To be sure, respectful actions are also partly defined by social convention, and people can consequently mimic outward signs of respect without harboring respectful attitudes, just as they can tell the truth without being honest. But the attitudes are in an important sense prior to the conventions and thus the actions: we first identify what attitudes are consistent with respect and only then ask how a person fluent in the relevant culture would express those attitudes through his actions.

To put this point another way, a different relationship obtains among reasons, actions, and attitudes in the case of respect than obtains among the three in the case of honesty. In the case of honesty, the value of the action of truth telling gives us a reason to

⁷ For the view that respect for persons is intrinsically good, see Joseph Raz, “On Respect, Authority, and Neutrality: A Response,” *Ethics* 120 (2010): 284.

⁸ Hence Aristotle’s distinction between the “self-controlled” person and the truly virtuous one (*Nicomachean Ethics*, bk. VII).

perform that action, and general principles of rationality tell us to perform the action for precisely that reason and thus to develop attitudes that will dispose us to do so. By contrast, in the case of respect, the value of the object of respect (e.g., a person) gives us a reason to have certain attitudes toward that object but does not tell us directly what actions to perform.⁹ Rather, only after we have identified the attitudes we should cultivate can we determine what actions are consistent or inconsistent with those attitudes—that is, what actions are respectful or disrespectful. Once again, respectful attitudes are definitionally prior to respectful actions.¹⁰

It is this definitional priority that necessitates a perfectionist account of antidiscrimination law. To condemn acting on discriminatory intentions as wrongful, we must first explain why such intentions are inconsistent with a properly respectful attitude. Although antiperfectionism may legitimately note when one takes another person's race or sex as a reason to treat her less favorably than one would treat people of other races or the other sex, it may not explicate the wrongfulness of such intentional action in terms of attitudes, which partly constitute moral character and thus implicate questions about the good life. To be sure, even on a respect-based account, antidiscrimination law does not directly assess individual discriminators' actual characters; it does not scrutinize the attitudes with which they act. But it nevertheless considers their actions to be wrong only because their intentions necessarily imply a morally defective character—namely, a character that lacks appropriately respectful attitudes toward other persons. To seek to

⁹ This statement must be qualified to account for the reasons that the value of an object gives us to preserve the object and refrain from destroying it. But our duties to preserve value and refrain from destroying it are not particularly relevant to the issue of discrimination.

¹⁰ This argument should not be taken to endorse the moral theory known as "virtue ethics." According to one version of that theory, our reasons for action "are *grounded* in the virtues" (Roger Crisp, "Modern Moral Philosophy and the Virtues," in *How Should One Live?*, ed. Roger Crisp (Oxford: Oxford University Press, 1996), 7). The argument in the text, by contrast, retains the priority of reason over virtue. It simply notes that, at least in the case of respect, reason sometimes cannot tell us what to do without first telling us what attitudes to cultivate.

combat a wrong that is necessarily defined in terms of moral character is an inescapably perfectionist purpose.

We might initially question the claim that one who acts on discriminatory intentions *necessarily* has a bad character. The example of the racially tolerant barber, considered in Chapter 5, heightens our doubts: even if the barber disrespects his friend and employee by capitulating to his clients' racism, he does not seem to be a vicious person on that count, but rather seems to display the economic prudence that we would expect of anyone in similar circumstances. To be sure, we would laud and admire the barber if he decided to resist his racist clients and sacrifice his economic self-interest for racial justice. But such moral courage and self-sacrifice would seem to be supererogatory, the act of a moral saint rather than a merely virtuous person.

On closer inspection, however, we discover that the barber does indeed display character flaws. Most obviously, he betrays his friend, the black employee whom he fires, by privileging his economic well-being over their friendship. Over all, the barber may be the consummate friend, but at least on this one occasion, he has not acted as a true friend would. To that extent, then, he lacks the virtue of friendship.

A similar analysis applies to the virtue of respect. The barber may very well be a generally respectful person, never discriminating based on race except when his back is to the wall, when his very economic survival depends on it. But so long as he discriminates even just once, he renders himself to that extent a disrespectful person, the kind of person who is willing to be disrespectful at least occasionally, even if not regularly.

This conclusion seems more plausible once we appreciate the distinction between *manifesting* a bad character and *constituting* one. As we have seen, we can understand a person's character as the set of settled dispositions that prompt him to act in certain ways, so that his actions manifest his character traits. On this Humean conception of character,

the barber is not a disrespectful person, since he lacks a settled disposition to discriminate against—and thereby disrespect—others. But we can also understand character as the product of one’s actions, so that each action one performs—however aberrant—partly constitutes one’s character. On this Aristotelian conception of character, the barber, though perhaps not a generally disrespectful person, is a partly disrespectful one, since he performed a disrespectful action at least once. To note this failing, of course, is not to condemn the barber as a vicious person *tout court*. Rather, it is to render the more nuanced judgment that, for all his merits, the barber has incurred a significant demerit that marks him as a less-than-fully-respectful person.

It might also be objected that antiperfectionism can actually attend to at least some of our attitudes without having to take a controversial position on the good life. So long as our account of respect and its concomitant attitudes is sufficiently thin—that is, sufficiently ecumenical among competing conceptions of the good life—we might think that antiperfectionism can authorize the government to prohibit disrespectful actions and even to foster respectful attitudes in its citizens. According to this objection, only because the previous chapter espoused an especially thick conception of respect for persons, one that took a definite position on the controversial question of the basis of the unique value of persons, does it seem that a respect-based account of the wrongfulness of discrimination necessitates a perfectionist account of antidiscrimination law.

A number of contemporary liberal philosophers have purported to develop thinner, antiperfectionist accounts of respect for persons. For example, Ronald Dworkin has argued that all persons are entitled to “equal concern and respect” and that this injunction *requires*, rather than precludes, governmental neutrality among competing conceptions of

the good.¹¹ John Rawls has similarly argued that persons are equal not in virtue of some “comprehensive” notion, such as Raz’s notion of ultimate value, but rather in virtue of the “political” notion of our ability “to be fully cooperating members of society.”¹² As we saw in Chapter 2, Rawls derives from this notion a full-blown theory of justice. More relevant for our present purposes, he also uses it to develop an account of respect for persons (though he does not use that precise term). In particular, Rawls argues that an antiperfectionist government “may still affirm the superiority of certain forms of moral character and encourage certain moral virtues”—namely, the “political virtues,” which include “the virtues of fair social cooperation such as the virtues of civility and tolerance, of reasonableness and the sense of fairness.”¹³ Inculcating these virtues, moreover, may legitimately involve “discouraging various kinds of religious and racial discrimination.”¹⁴ If an antiperfectionist government may directly inculcate certain character traits in its citizens, then, a fortiori, it may define wrongs like discrimination in terms of those traits.

Both Dworkin’s and Rawls’s thin theories of respect, however, are vulnerable to two criticisms. First, it is far from clear that either theory remains truly antiperfectionist. Indeed, Dworkin concedes that we must invoke a conception of the good life in order to cash out his injunction to treat people with “equal concern and respect” (though once again, he believes that a true conception of the good life will, somewhat paradoxically, support governmental neutrality among competing conceptions of the good life).¹⁵ Similarly, despite Rawls’s repeated protestations to the contrary, the “political virtues” do indeed seem to presuppose a conception of the good life. For the development of the “political virtues” will inevitably hinder the development of other, nonpolitical virtues,

¹¹ See, e.g., Ronald Dworkin, *A Matter of Principle* (Cambridge, Mass.: Harvard University Press, 1985), ch. 8.

¹² Rawls, *Political Liberalism*, 19.

¹³ *Ibid.*, 194.

¹⁴ *Ibid.*, 195.

¹⁵ See Ronald Dworkin, *Sovereign Virtue* (Cambridge, Mass.: Harvard University Press, 2000), ch. 6.

and some conceptions of the good will prioritize the latter virtues over the former ones. Rawls would, of course, respond that such conceptions neglect the priority of the right over the good, or the “political” over the “comprehensive,” a fundamental requirement of justice. But other than affixing such conceptions with the question-begging label “unreasonable,” Rawls gives us no neutral (i.e., non-comprehensive) reason to reject their ranking of virtues.¹⁶

Second, even if Dworkin’s and Rawls’s theories of respect are genuinely antiperfectionist, both appear to be too thin to underwrite the kinds of duties enforced by antidiscrimination law. The problem is more acute for Dworkin’s theory, whose injunction to treat every person with “equal concern and respect” binds only the government, not private persons. It is one thing to enjoin the government to respect persons by, *inter alia*, not discriminating against them based on race or sex, but it is quite another to enjoin persons themselves to do so,¹⁷ and Dworkin gives us little guidance about how to move from the political to the personal without violating the strictures of antiperfectionism. Even Rawls’s theory, with its unequivocally personal “political virtues,” faces a similar problem. For Rawls dilutes those virtues in order to reconcile them with the rest of his theory. For example, after asserting that the government may foster the virtue of tolerance by “discouraging” discrimination, Rawls adds the caveat: “in ways consistent with liberty of conscience and freedom of speech.”¹⁸ As we saw in Chapter 2, the government’s efforts to combat discrimination will often conflict with basic rights and liberties, as Rawls understands them. It is difficult to see how we can render those efforts more fruitful simply by recasting them as measures for inculcating “political virtues”

¹⁶ Cf. Joseph Chan, “Legitimacy, Unanimity, and Perfectionism,” *Philosophy & Public Affairs* 29 (2000): 5-42; Stephen Mulhall and Adam Swift, *Liberals and Communitarians*, 2nd ed. (Oxford: Blackwell Publishing, 1996), ch. 7.

¹⁷ Cf. John Gardner, “Liberals and Unlawful Discrimination,” *Oxford Journal of Legal Studies* 9 (1989): 2-3.

¹⁸ Rawls, *Political Liberalism*, 195.

rather than for achieving socioeconomic equality. Although these arguments do not conclusively rule out an antiperfectionist account of respect-based duties that is sufficiently robust to justify the full extent of antidiscrimination law, they do cast serious doubt on the viability of such an account.

III. RESPECT, RIGHTS, AND JUSTICE

This thesis has attempted to construct a dilemma for contemporary liberalism. In Part I of this thesis, we saw how three of the most prominent contemporary liberal doctrines of political legitimacy fail to justify the full extent of antidiscrimination law. In each case, this failure was attributed to a common source: each doctrine's focus on the effects discrimination tends to have for its victims. Because these effects do not always materialize, an effects-based account of the wrongfulness of discrimination will inevitably end up condoning certain instances of the conduct. In light of the failure of effects-based accounts to justify all applications of antidiscrimination law, the previous chapter sought to identify an intrinsic feature of discrimination that renders every instance of the conduct wrongful—namely, the incompatibility of acting on discriminatory intentions and having an appropriately respectful attitude toward the victims of discrimination. Finally, the previous section of this chapter argued that a respect-based account of the wrongfulness of discrimination necessitates a perfectionist account of antidiscrimination law, since the wrongfulness of acting on discriminatory intentions cannot be specified independently of discriminators' moral characters. The dilemma for contemporary liberalism is now fully apparent: either contemporary liberalism will continue to espouse an effects-based account of the wrongfulness of discrimination, in which case it will contravene our intuition that all instances of discrimination are wrong and that all the central applications of antidiscrimination law are justified, or it will

espouse a respect-based account, in which case it will contravene its own antiperfectionism.

Given the significant costs attending each option, many liberals will now wish to argue that the dilemma is a false one. In particular, they will wish to deny the second horn of the dilemma, insisting that an account of the wrongfulness of discrimination that looks beyond the conduct's effects for its victims need not attend to discriminators' moral characters and thus need not entail a perfectionist account of antidiscrimination law. For it seems that discrimination necessarily violates its victims' *rights* and is therefore always *unjust*. To be sure, it would be question-begging simply to assert the existence of a right not to be discriminated against. But the previous chapter suggested the most plausible basis for such a right: the respect to which we all are entitled simply in virtue of our status as persons. If we can specify a right to respect—and its correlative duties—independently of people's moral characters, then it seems we need not espouse perfectionism in order to justify antidiscrimination law after all.

The notion of a right to respect, however, confronts an immediate challenge, for we might doubt whether there can even be such a right—or at least the kind of right that can ground a further right not to be discriminated against. For example, although Raz believes that we have reasons to have certain respectful attitudes toward others, he denies that others have a right for us actually to have those attitudes.¹⁹ Of course, Raz must believe that there are at least some respect-based rights. In particular, it seems that the reasons we have not to destroy other persons correspond to a right in other persons not to be destroyed. Given his apparent willingness to recognize at least some respect-based rights, we might wonder why Raz refuses to recognize a respect-based right for other people to have respectful attitudes toward us.

¹⁹ Raz, "On Respect," 282.

The most plausible answer is that Raz believes that we can have no right for others to cultivate certain attitudes because others cannot be under a duty to do so. According to Raz, a person has a right if and only if “an aspect of [his] well-being (his interest) is a sufficient reason for holding some other person(s) to be under a duty.”²⁰ The existence of a right thus “depends not only on the interest of the right-holder but, at the very least, also on that of the person who is to be subject to that duty.”²¹ Raz seems to believe that a duty to cultivate respectful attitudes toward others would simply be too demanding for the bearer of the duty and, consequently, that any interest we might have in others’ cultivating respectful attitudes toward us is insufficient to ground such a duty.²² Indeed, many moral philosophers believe that complying with such a duty would prove not just overly demanding, but impossible, since we cannot choose our attitudes.²³ But this whole line of argument is something of a red herring. For antidiscrimination law requires us not to cultivate respectful attitudes (which would be impossible), but rather to forbear acting with certain intentions that preclude such attitudes. Even if we cannot choose to have respectful attitudes, we can recognize when our acting on a particular reason would be incompatible with our harboring such attitudes toward some other person and refrain from so acting. Because a duty to display such awareness and restraint seems to follow from people’s interest in being treated in accordance with the ultimate value they possess, it does seem that there can be the kind of right to respect that would ground a further right not to be discriminated against.²⁴

²⁰ Joseph Raz, *The Morality of Freedom* (Oxford: Oxford University Press, 1986), 166.

²¹ Joseph Raz, *Ethics in the Public Domain* (Oxford: Oxford University Press, 1994), 35.

²² Note, however, that Raz believes that parents have a duty to love their children and that he conceives of love as an “attitude” (though he also believes that the duty to love does not apply much beyond the parent/child relationship) (*ibid.*, 10-13).

²³ See, e.g., David Ross, *The Right and the Good*, ed. Philip Stratton-Lake (Oxford: Oxford University Press, 2002), 4-5.

²⁴ This argument has certain affinities with Harry Frankfurt’s account of free will. See Harry G. Frankfurt, *The Importance of What We Care About* (Cambridge: Cambridge University Press, 1988), ch. 2.

This kind of right to respect becomes even more plausible if, like many political theories, we associate rights with justice. Following Aristotle and Aquinas, John Finnis identifies a notion of justice that concerns all “relationships and dealings between persons . . . in which there can be question of what is fitting, fair, or just as between the parties to the relationship.”²⁵ At least the kind of disrespect involved in discrimination violates justice in this sense. As we discovered in the previous chapter, the wrong of discrimination consists in taking another person’s race or sex as a reason to treat her less favorably than one would treat people of other races or the other sex. Acting on such an intention is disrespectful because it is inconsistent with an attitude that regards the victim as a being of ultimate value, a being whose life stands to go better or worse depending on her engagement with value. Such disrespect is also unjust because it violates the precept that “the human person must seek to realize and respect human goods not merely in himself and for his own sake but also in common, in community” with others—the very precept that grounds the notion of justice identified above.²⁶ If it is unjust to treat someone as if he enjoyed less than a full claim to engagement with value, then it seems natural to suppose that such treatment also violates his rights.

The more urgent question for our purposes, however, is not whether there is a right to respect, but rather, if there is such a right, whether it accords with antiperfectionism. The most plausible basis for an antiperfectionist account of respect-based rights is found in Kantian political philosophy. Kant famously distinguishes between two kinds of duties: “duties of right” and “duties of virtue.”²⁷ According to Arthur Ripstein’s recent reconstruction of Kant’s political philosophy, the difference between these two kinds of duties is “[t]he difference between inner willing and outer

²⁵ John Finnis, *Natural Law and Natural Rights* (Oxford: Oxford University Press, 1980), 177-78. Cf. John Finnis, *Aquinas* (Oxford: Oxford University Press, 1998), 133.

²⁶ Finnis, *Natural Law*, 161.

²⁷ See generally Immanuel Kant, *The Metaphysics of Morals*, trans. and ed. Mary Gregor (Cambridge: Cambridge University Press, 1996).

freedom.”²⁸ More specifically, whereas duties of virtue concern “the maxim on which an action is done,” the reason for which it is performed, duties of right concern “the outer form of interaction between persons.”²⁹ So long as our external actions conform to the proper standards governing interpersonal conduct—in particular, so long as we avoid violating others’ rights—we fulfill our duties of right. By contrast, to fulfill our duties of virtue, we must not only comport ourselves appropriately toward others, but also act from a moral “maxim”—that is, take some moral reason as our reason for action.³⁰

If duties of right simply reduced to duties of virtue, then the government could not enforce the former without enforcing the latter, and Kantian political philosophy would necessarily be perfectionist. But in fact, Kant believes that the two kinds of duties can, and often do, come apart. In particular, fulfilling our duties of virtue does not guarantee that we will also fulfill our duties of right. For we are embodied beings who cannot simultaneously occupy the same physical space and whose “deeds can [thus] come into conflict regardless of the maxims on which we act.”³¹ One can, for example, trespass on another person’s land innocently, acting on a perfectly laudable “maxim.”³² Kant thinks that this kind of scenario bespeaks a need for two dimensions of moral assessment: the domain of virtue and the domain of right, the domain of the reasons for which we act and the domain of the relations we form, the domain of what we do simpliciter and the domain of what we do to others.³³ Because they are “relational,” concerning only what we do to others rather than the reasons for which we do it, duties of right remain wholly distinct from duties of virtue.³⁴

²⁸ Ripstein, *Force*, 358.

²⁹ *Ibid.*, 11.

³⁰ *Ibid.*, 374.

³¹ *Ibid.*, 376.

³² *Ibid.*, 380-81.

³³ *Ibid.* Cf. Scanlon, *Moral Dimensions*, 61 n.15. Kant explores the domain of virtue in his *Groundwork of the Metaphysics of Morals*, trans. and ed. Mary Gregor (Cambridge: Cambridge University Press, 1997).

³⁴ Ripstein, *Force*, 109.

So long as it enforces only duties of right and not duties of virtue, the government can comply with antiperfectionism, since duties of right do not concern the reasons for which people act, the attitudes they adopt, or, therefore, the moral characters they cultivate. And as many contemporary liberal political philosophers suppose, Kant preserves the antiperfectionism of his political theory by insisting that only duties of right are legitimately enforceable.³⁵ Indeed, Kant believes that it is *impossible* to enforce duties of virtue, since one cannot be coerced to adopt a particular “maxim,” to act for a particular reason.³⁶ Focused on actions, duties, and rights rather than reasons, attitudes, and characters, Kant’s political theory is highly congenial to antiperfectionism.

This rendition of Kantian political philosophy, however, seems to preclude a legally enforceable duty of (and thus right to) respect, for such a duty appears to be a duty of virtue rather than a duty of right, concerning the reasons for and attitudes with which one acts rather than one’s treatment of others alone. We thus might wonder how Kantian political theory would justify a right not to be discriminated against and a correlative duty not to discriminate. According to Ripstein, such a right would be grounded in the basic concept underlying Kant’s entire political theory: “independence.” All persons, Kant believes, enjoy a fundamental right to independence, which consists in being one’s “own master”—that is, in “not being subordinated to the choice of any *other* particular person.”³⁷ One enjoys independence, in other words, so long as one can set and pursue one’s own purposes rather than being made a mere means to serve the purposes of others.³⁸

³⁵ See, e.g., *ibid.*, 14, 160. See also Hamish Stewart, “The Limits of the Harm Principle,” *Criminal Law and Philosophy* 4 (2010): 30-31.

³⁶ Ripstein, *Force*, 269. See also Korsgaard, *Creating*, 19-20.

³⁷ Ripstein, *Force*, 4.

³⁸ *Ibid.*, 34. Thus stated, the Kantian notion of “independence” strongly resembles the republican conception of freedom as “non-domination.” See Philip Pettit, *Republicanism* (Oxford: Oxford University Press, 1997), chs. 2 and 3.

Discrimination appears to compromise its victims' independence, thus understood. More specifically, Ripstein believes that the duty not to discriminate follows from the kind of logic that justifies the more general Kantian duty to support the poor within a system of private property. Although, according to Ripstein, Kantian principles generally countenance socioeconomic inequality stemming from a system of private property, such principles condemn inequality that is severe enough to compromise people's independence and support an enforceable duty of right to help to offset such inequality through contributions to a program of redistributive taxation.³⁹ If the state itself thus may not "set up a system of property that would allow one person to become fully subject to the choice of another," then neither may it "set[] up a system of private property that gave private persons the unrestricted power to use their private property to establish one."⁴⁰ But that, we might think, is precisely the kind of system of private property the government would establish if it failed to prohibit private employers and purveyors of public accommodations from discriminating based on race or sex. On this view, antidiscrimination law guarantees people's right to independence by protecting them from becoming subject to the wills of employers and purveyors of public accommodations.

This independence-based account of antidiscrimination law, however, suffers from at least two shortcomings. First, it is unclear whether the account is actually antiperfectionist, for some Kantian scholars argue that redistributive taxation enforces a duty of virtue rather than a duty of right.⁴¹ If these scholars are correct, then, on Ripstein's account, the duty not to discriminate is also a duty of virtue, one the government may not enforce without espousing perfectionism.

Second, and more fundamentally, an independence-based account of antidiscrimination law confronts the same difficulty that thwarted the harm-based

³⁹ See generally Ripstein, *Force*, ch. 9.

⁴⁰ *Ibid.*, 292.

⁴¹ See, e.g., Onora O'Neill, *Constructions of Reason* (Cambridge: Cambridge University Press, 1989), 231.

accounts considered in Chapters 3 and 4—namely, an inability to justify the prohibition of instances of discrimination that have no deleterious consequences for their victims. To be sure, unlike harm-based accounts, an independence-based account does not consider the concrete effects discrimination has for the interests of its victims. As the case of the so-called happy slave illustrates, a person’s independence can be compromised—he can be made subject to the will of some other person—even if he remains unharmed.⁴² But *pace* Ripstein, an independence-based account of discrimination remains focused on the conduct’s consequences for its victims—namely, the conduct’s erosion of its victims’ independence. An independence-based account of discrimination will thus justify antidiscrimination law only insofar as the conduct actually has such an impact.

In fact, it seems unlikely that discrimination will always compromise the independence of its victims. For even if discrimination prevents its victims from pursuing *some* of their purposes, it rarely will prevent them from pursuing *all* their purposes. The victims will not, in other words, always be rendered “*fully* subject to the choice of another” (i.e., the discriminatory employer or purveyor of a public accommodation). A victim might, for instance, have access to job opportunities comparable to the one he was denied. In such a case, the victim will be able to pursue many, if not all, of his purposes independently of the discriminator’s will. If the government prohibited discrimination even when the victim enjoyed this degree of freedom, it would seem unduly to compromise the discriminator’s own independence, constraining his freedom in order to further the specific purposes of others.⁴³ It thus appears that an independence-based account of discrimination can justify antidiscrimination law only when discrimination is

⁴² Ripstein, *Force*, 36, 269.

⁴³ For an analogous critique of Ripstein’s argument for redistributive taxation, see Victor Tadros, “Independence Without Interests?,” *Oxford Journal of Legal Studies* 31 (2011): 205. Recall that, on Ripstein’s reading of Kant, duties of right are distinct from duties of virtue. Thus, for purposes of the law, a person’s freedom is limited not by the moral quality of his actions, but only by the freedom of others.

relatively entrenched and pervasive, compromising its victims' independence enough to offset antidiscrimination law's interference with the freedom of discriminators.⁴⁴

A Kantian right to independence thus cannot ground a right not to be discriminated against—at least not a nondiscrimination right that is sufficiently robust to justify all applications of antidiscrimination law. This failure might prompt us to reconsider a right to respect as the basis for a right not to be discriminated against. Recognizing a right to respect would indeed justify all applications of antidiscrimination law, since, as the previous chapter argued, every instance of discrimination necessarily disrespects its victim. But this victory would be a Pyrrhic one for contemporary liberalism. For we cannot recognize a right to respect without defining the right's correlative duties, and as we saw in the previous section, we cannot define duties of respect without attending to people's reasons, attitudes, and moral characters. To put the point in Kantian terms, duties of respect are not duties of right, but duties of virtue, the kinds of duties the government may enforce only by espousing perfectionism. We thus cannot reconcile antidiscrimination law with antiperfectionism simply by recasting the wrong of discrimination in terms of rights and justice.

This conclusion highlights an important feature of the right not to be discriminated against, understood as a right to a kind of respect, and the injustice of discrimination, understood as a kind of disrespect. Some perfectionists attempt to link justice and moral character by defining justice itself as one of the virtues. Indeed, Aristotle goes so far as to posit a "general" sense of justice that encompasses "complete excellence [virtue] . . . in

⁴⁴ Ripstein could, of course, modify his account of independence to fill these gaps. For example, he might argue that, even when discrimination does not *actually* interfere with its victims' independence, it still creates a *risk* of such interference, thereby endangering the victims' ability to pursue their own purposes independently of other people's wills (cf. Victor Tadros, "Harm, Sovereignty, and Prohibition," *Legal Theory* 17 (2011): 60). The critique of this rejoinder would parallel the critique of similar efforts to modify the harm principle in Chapter 3.

relation to another person.”⁴⁵ Other theorists tend toward the opposite extreme, seeking to dissociate the concept of justice from moral character altogether. In a Kantian vein, such theorists insist on a rigid distinction between unjust actions and immoral characters, between the quality of our treatment of other people and the quality of the motivations that prompt that treatment.⁴⁶ Both positions are overdrawn. On the one hand, there are some injustices, like distributive injustices, that can arise from myriad vicious—and perhaps even some virtuous—dispositions and that can consequently be specified wholly independently of people’s moral characters. On the other hand, there are other injustices that supervene on particular character traits, defying the dichotomy posited by contemporary liberals between justice and character.⁴⁷

The form of disrespect wrought by discrimination is one such injustice. Unlike other injustices, the wrong of discrimination is defined in terms of one specific character failing—namely, the lack of a properly respectful attitude entailed by one’s taking another person’s race or sex as a reason to treat her less favorably than one would treat people of other races or the other sex. To be sure, this failing can stem from myriad further vices, such as racism, avarice, or cowardice. But regardless of its precise provenance, the fault of disrespect is a distinct failing that both further blemishes the characters of the people who exhibit it and, when acted on, engenders its own, unique injustice. Only by considering the fault of disrespect can we explain how acting on a discriminatory intention transmutes the mere denial of some opportunity into a genuine injustice. Therefore, even if we understand antidiscrimination law as a measure for protecting rights and securing justice, we still must be willing to espouse perfectionism in order to justify all the legislation’s applications.

IV. RESPECT, PERFECTIONISM, AND THE PRUDENTIAL LIMITS OF THE LAW

⁴⁵ Aristotle, *Nicomachean Ethics*, 1129b25-30.

⁴⁶ See, e.g., Bernard Williams, *Moral Luck* (Cambridge: Cambridge University Press, 1981), 84, 86.

⁴⁷ Cf. *ibid.*, 86-87 (tentatively acknowledging the existence of such injustices).

But just how broadly does a respect-based perfectionism sweep? In adopting a doctrine of political legitimacy that authorizes the government to prohibit discrimination, just how much other legal moralism do we countenance? We might worry that such a doctrine comprehends forms of behavior that strongly implicate the autonomy of the people who engage in it and that consequently merit toleration. For example, if the government may proscribe discrimination in employment and the provision of public accommodations, why may it not do so in other, less public settings as well? After all, discriminators treat their victims no less disrespectfully merely by confining their discriminatory conduct to putatively private domains like dinner parties and club meetings; to discriminate based on race or sex is always to disrespect one's victims, regardless of where, exactly, the discrimination occurs. It also appears that a perfectionist, respect-promoting doctrine of political legitimacy will license the government to obtrude into all domains—private as well as public—for the sake of curtailing not just discrimination based on race or sex, but other forms of disrespect as well. A respect-based perfectionism seems to end up permitting the government to regulate people's activities too intrusively and to restrict their autonomy too severely.

Any proponent of perfectionism must acknowledge the considerable truth in this line of criticism. In particular, no perfectionist, respect-promoting doctrine of political legitimacy can ground a *principled* objection to the kinds of regulation considered in the previous paragraph. Such a doctrine, in other words, cannot condemn *as unjust* attempts by the government to curb discrimination based on race or sex in indisputably private domains or other forms of disrespect in any setting. To pronounce a given piece of regulation just in principle, however, is not to deem it permissible in practice. For the ultimate propriety of any governmental action depends on its compliance not only with

principle, but also with prudence.⁴⁸ Just like their contemporary liberal counterparts, all nonfundamentalist perfectionist doctrines of political legitimacy recognize prudential limits on governmental power, countenancing only regulation that is efficacious as well as just.⁴⁹ We thus should not infer that perfectionism sanctions the extension of antidiscrimination law into the private realm or the prohibition of all disrespectful conduct merely from the fact that it fails to condemn those measures as unjust. Rather, we must first ensure that they achieve their professed aims efficiently, that they successfully curtail their targeted conduct without infringing other important values.

We have good reason to suppose that, by insinuating antidiscrimination law into indisputably private domains, the government would indeed pay a considerable price—too great a price to warrant any moralistic gains that such an intrusion might realize. The government must dedicate its limited regulatory resources to remedying only a relatively few immoralities, abiding other ones neither because they are any less immoral nor because their perpetrators are any less vicious, but because they tend to be less harmful for their victims. Were the government to attempt to forestall every instance of discrimination in every domain, it would end up spreading itself too thin and actually prevent fewer discriminatory acts than it would if it focused on only a subset of the conduct. The government is thus justified in adopting the very strategy employed by modern antidiscrimination regimes: regulating those settings, such as employment, where discrimination is most likely to harm its victims while withdrawing from those, like dinner parties, where the conduct proves less deleterious.

In making this prudential calculation, the government might initially appear to be attempting to comply with the dictates of the liberal harm principle. But any such resemblance between perfectionist and liberal doctrines of political legitimacy is in fact

⁴⁸ See generally Robert P. George, *Making Men Moral* (Oxford: Oxford University Press, 1993).

⁴⁹ See generally George Sher, *Beyond Neutrality* (Cambridge: Cambridge University Press, 1997), ch. 5.

specious. As we have seen, whereas the harm principle categorically precludes the prohibition of harmless conduct, a perfectionist doctrine of political legitimacy actually authorizes such regulation (at least in principle); only because the government possesses limited resources does perfectionism require it to heed a prudential imperative and prioritize the regulation of harmful instantiations of various immoralities over the regulation of innocuous ones.

Even if the government had unlimited resources, it still would likely confront other significant practical impediments to extending antidiscrimination law's purview into more private domains. In fact, it seems that such an expansion would prove self-defeating, ultimately failing even to uncover—let alone curtail—the very conduct it purported to regulate. For while we can often determine the grounds on which an employer decides whether to hire a job applicant, we can rarely have such insight into, say, one person's decision whether to befriend another. Nor would the government experience any more success were it to attempt to scrutinize its citizens' fraternal choices as part of its antidiscrimination program. On the contrary, the government seems particularly unsuited to discover subtly discriminatory social practices, an incapacity that militates against extending antidiscrimination law to the settings where we are most likely to encounter such practices.

Just as prudential considerations militate against extending antidiscrimination law into private domains, so, too, do they preclude regulating forms of disrespect other than discrimination based on race or sex—even in the most indisputably public settings. For the government could rarely uncover forms of disrespect based on less perceptible traits. For example, to determine whether a particular employer has discriminated against a job applicant based on social sophistication, the government must not only identify a class of people whom others tend to regard as inferior beings lacking ultimate value simply

because of their relative provinciality, but also gauge the employer's perceptions of and attitudes toward that trait in general and the applicant's embodiment of it in particular. The government seems incapable of making such discerning assessments. Moreover, unlike discrimination based on race or sex, discrimination based on social sophistication remains sufficiently anomalous to leave most of its potential victims unscathed and thus to constitute the kind of immorality that the government can abide with relative equanimity. Once again, we must be careful not to mistake this prudential limitation on the regulation of other kinds of disrespectful discrimination for a principled one. But prudential limitations nonetheless significantly circumscribe governmental power, rendering perfectionism a much less intrusive ideology than we might initially worry.

V. RESPECT, PERFECTIONISM, AND THE PRINCIPLED LIMITS OF THE LAW

The previous section will have allayed few of the fears that perfectionism tends to provoke among contemporary liberals. For while it recognized certain prudential reasons to distinguish among different kinds of disrespectful conduct, it recognized no *principled* reasons to distinguish disrespectful conduct from, say, purely self-regarding immoralities. And yet, we tend to think that these latter immoralities should be immune to governmental regulation absolutely, for reasons of principle, not merely contingently, for reasons of prudence. It is not enough to acquit a perfectionist doctrine of political legitimacy of the charge of fundamentalism, of a willingness to prohibit immorality at any cost; we must also show the doctrine to be tolerant, willing to recognize at least some principled limits on the government's power to prohibit immorality. The challenge, however, is to identify such principled limits without condemning certain applications of antidiscrimination law, a challenge that the contemporary liberal doctrines of political legitimacy considered in Part I of this thesis could not meet.

Those versions of perfectionism that do recognize principled, rather than merely prudential, limits to the law seem equally unable to meet this challenge. For example, in interpreting Thomas Aquinas's political philosophy, John Finnis propounds a perfectionist theory that, while explicitly founded on a conception of the good life, nevertheless develops a distinctly political account of the "common good" of society. In particular, the theory defines the common good as "a set of conditions which enables the members of a community to attain for themselves reasonable objectives, or to realize reasonably for themselves the value(s), for the sake of which they have reason to collaborate with each other (positively and/or negatively) in a community."⁵⁰ According to this definition, the common good comprises not the actual flourishing of society's individual members, but rather the social prerequisites for its achievement. The government thus bears responsibility for only a truly public good that is both "limited" and "instrumental"—"limited" because it restricts its purview to the social milieu and shared goods conducive to individual human flourishing; "instrumental" because it seeks not to supplant, but to facilitate individual human flourishing.⁵¹

We see just how limited the political common good actually is when we consider its precise content. According to Finnis, the political common good comprehends only those "goods (and virtues) which are intrinsically interpersonal, other-directed . . . , person to person"—namely, "justice and peace."⁵² The justice comprehended by the political common good, moreover, is not "complete justice, which true virtue requires of us," but merely "the peaceful condition needed to get the benefits . . . of social life and avoid the burdens of contention."⁵³ In other words, to realize the political common good,

⁵⁰ Finnis, *Aquinas*, 155.

⁵¹ *Ibid.*, 235-39.

⁵² *Ibid.*, 226-27.

⁵³ *Ibid.*, 227-28.

the government need establish only that degree of communal concord that must obtain if people are to pursue their valuable projects and activities with assurance and in security.

If, as most political theories contend, the government must restrict itself to establishing and maintaining the political common good, then to determine the content of the political common good is necessarily to delineate the boundaries of legitimate political power. Because it seeks to secure the social preconditions for human flourishing, rather than human flourishing itself, Finnis's Thomistic account of the common good forbids the government to cultivate its citizens' virtue directly, as a means of ensuring their adherence to the good life.⁵⁴ Finnis adduces two reasons for this limitation—one of prudence and one of principle. First, as a matter of sheer practicality, the government cannot discern “the furthest and deepest interests and dispositions underlying [people's actions],” which, in turn, constitute the very moral characters perfectionist governments purport to cultivate.⁵⁵ Second, as a matter of basic justice, the government violates its citizens' rights whenever it restricts personal behavior that has no effect on “public well-being,” that is in no way “stamped with a public character.”⁵⁶ Both self-defeating and unjust, coercive moral paternalism apparently exceeds the limits of the political common good.

Finnis's ostensibly liberal account of the common good, however, is in tension with other, more perfectionist aspects of his theory. Indeed, it seems that his theory founders on the horns of a dilemma. On the one hand, if we adopt Finnis's limited conception of the common good, forbidding the government to cultivate its citizens' characters except insofar as such moralism is necessary to secure basic justice and peace, we once again call the legitimacy of antidiscrimination law into question. For at least in relatively tolerant social environments, discrimination need not imperil justice and peace

⁵⁴ Ibid., 228.

⁵⁵ Ibid., 241.

⁵⁶ Ibid., 241-42.

enough to hinder anyone's human flourishing—either that of the members of society in general, who will likely remain uncorrupted by isolated instances of discrimination, or that of the victims of discrimination in particular, who will usually continue to enjoy a range of opportunities sufficient for a good life. On one reading, then, Finnis's theory reduces to just another version of liberal perfectionism, permitting the government to endorse a conception of the good life yet enjoining it from implementing many of the moralistic regulations required for a fully effective antidiscrimination regime.

Alternatively, we can construe Finnis's theory as a rather robust version of perfectionism that countenances rather intrusive moralistic regulation. Such an interpretation gains significant plausibility from Finnis's own statements regarding the relationship between public justice and peace, on the one hand, and personal virtue, on the other. For example, although Finnis initially identifies the political common good with "peace," he ultimately defines the latter concept so capaciously that it embraces "not only concord (absence of dissension, especially on fundamentals) and willing agreement between one person or group and another, but also harmony . . . amongst each individual's own desires."⁵⁷ Finnis also explicitly acknowledges the need to inculcate private virtues as well as public ones to foster a social environment conducive to human flourishing.⁵⁸ In some of his other work, Finnis takes this moralism even further, advocating "the use of coercive measures" simply for the paternalistic purpose of "dissuad[ing] [people] from morally evil forms of life."⁵⁹ Such unadulterated perfectionism appears to contradict Finnis's earlier protestations that his theory balks at the most intrusively moralistic regulations; hence the dilemma: either Finnis's theory defines the common good narrowly and thus precludes most forms of character

⁵⁷ *Ibid.*, 227.

⁵⁸ See, e.g., *ibid.*, 232.

⁵⁹ John Finnis, "Legal Enforcement of 'Duties to Oneself': Kant v. Neo-Kantians," *Columbia Law Review* 87 (1987): 456.

cultivation, in which case it suffers from many of the same deficiencies that plagued liberal perfectionism, or it underwrites a more moralistic agenda and authorizes the government to inculcate virtue, in which case it collapses into a thoroughgoing perfectionism—the very kind of perfectionism that Finnis purports to eschew.

If we are to legitimate the full extent of antidiscrimination law, then we must espouse a brand of perfectionism that hews more closely to the second, comprehensively perfectionist interpretation of Finnis’s theory than to the first, tentatively liberal one, a version that licenses at least some character-cultivating regulation not just for some ulterior purpose, but also for its own sake. For only such regulation can reach those instances of discrimination that do not significantly jeopardize justice and peace. More specifically, where discrimination is sufficiently anomalous to leave its victims relatively unscathed, the government has recourse to but one justification for proscribing every instance of the conduct: to forestall certain expressions of disrespect. To be sure, assuming everyone enjoys a right not to be treated with disrespect, the government necessarily promotes a kind of justice whenever it checks disrespectful conduct. But unlike the limited kind of justice Finnis associates with the political common good, this one transcends the basic prerequisites of the good life, comprehending facets of our moral characters and thus constituting a set of virtues to be acquired as well as a set of rights to be protected. Antidiscrimination law presupposes a thoroughgoing perfectionism.

Such a moralistic political theory, however, need not jettison all principled limits to the law. In particular, to legitimize the kind of moralism practiced by antidiscrimination law, we need not countenance the proscription of genuinely private vices, vices whose expression in no way affects anyone besides the vicious themselves. It is precisely such vices that Finnis attempts to insulate from governmental regulation; hence his repudiation of those perfectionist theories that would “mak[e] even secret and

truly consensual adult acts of vice a punishable offence against the state's laws.”⁶⁰ Robert George propounds one such theory, maintaining that “there is nothing in principle unjust about the legal enforcement of morals or the punishment of those who commit moral offenses”—private or otherwise.⁶¹ So long as there are no “prudential” obstacles to regulating a given vice, George contends, the government may legitimately attempt to curb it.⁶² As we now can see, a version of perfectionism like George's legitimates regulation that is much more intrusively moralistic than antidiscrimination law. For even if the fault of disrespect, as evinced through discrimination based on race or sex, does not imperil public justice and peace enough to hinder human flourishing, it undoubtedly affects people other than those who evince it—namely, the victims of discrimination, whose rights discriminators necessarily infringe. Therefore, to legitimize antidiscrimination law, we need espouse an account of the common good and a corresponding doctrine of political legitimacy that extend only slightly further than Finnis's, encompassing vicious conduct that affects people other than the vicious themselves but that nevertheless tends not to compromise public justice and peace. In subjecting such conduct to governmental regulation, moreover, we can continue to hold, with Finnis and against George, that the government does indeed commit a serious injustice when it attempts to curb genuinely private expressions of vice, that any plausible perfectionist theory of political legitimacy will impose on the government limits not merely of prudence, but also of principle. To sanction forms of legal moralism like antidiscrimination law is by no means to condone the most intrusive forms of coercive moral paternalism.

⁶⁰ John Finnis, “Is Natural Law Theory Compatible with Limited Government?,” in *Natural Law, Liberalism, and Morality*, ed. Robert P. George (Oxford: Oxford University Press, 1996), 8.

⁶¹ George, *Making Men Moral*, 1.

⁶² *Ibid.*, 2.

Lest we be tempted to infer from the relative moderation of this version of perfectionism any more significant affinity with contemporary liberalism, we should recognize the precise ways in which the former theory contravenes the fundamental tenets of the latter. *Pace* the antiperfectionist strands in contemporary liberalism, the version of perfectionism endorsed in the preceding paragraphs follows all other versions of perfectionism—including self-avowedly liberal ones—in unequivocally repudiating liberal public reason and endorsing a conception of the good life, which, in turn, it authorizes the government to promote. More radically, it departs from even liberal perfectionism in at least occasionally subordinating personal autonomy to other, perfectionist values—more specifically, in at least occasionally restricting people’s conduct simply to cultivate their characters. To be sure, no version of perfectionism presumes to engender virtuous conduct through such legal moralism alone. For virtue cannot be compelled; to act virtuously, one must deliberately choose for oneself to perform virtuous actions for virtuous reasons, a requirement that all plausible versions of perfectionism acknowledge.⁶³ But unlike liberal perfectionism, a more thoroughgoing perfectionism recognizes that the government can sometimes indirectly favor virtue by directly disfavoring vice and thus permits the government to infringe people’s autonomy to prevent them from performing certain vicious actions. It is precisely such oblique legal moralism that is practiced by antidiscrimination law and that distinguishes the legislation as a comprehensively perfectionist—though by no means morally fundamentalist—measure.

VI. CONCLUSION

How can antidiscrimination law possibly be perfectionist? After all, discrimination *wrongs* its victims, and if antiperfectionism can justify any governmental

⁶³ See, e.g., Finnis, “Legal Enforcement,” 456; George, *Making Men Moral*, 42-47.

regulation at all, it surely can justify regulation that prevents people from wronging others. Paradigmatically perfectionist regulation, by contrast, proscribes self-regarding immoralities, immoralities that blemish the moral characters of the people who engage in them but do not affect the interests or rights of others. But as this chapter has shown, we cannot reconcile antidiscrimination law with antiperfectionism simply by pronouncing discrimination to be a wrong to its victims. For the peculiar wrong of discrimination is defined in terms of the intentions of discriminators, the reasons for which they act. The wrongfulness of acting on such intentions, in turn, is defined in terms of the lack of appropriately respectful attitudes they entail. Because our attitudes partly constitute our moral characters, any prohibition of the wrong of discrimination will necessarily be perfectionist. We cannot purge antidiscrimination law of this perfectionist taint simply by positing a right to respect and conceiving of disrespect as an injustice, since the correlative duties of respect must still be explicated in terms of the duty bearers' moral characters.

Not only does antidiscrimination law have perfectionist underpinnings; it also has perfectionist effects, indirectly shaping the moral characters of actual and would-be discriminators through a kind of Aristotelian habituation. This moralism, however, need not entail moral fundamentalism. On the contrary, a respect-based perfectionist doctrine of political legitimacy can recognize prudential reasons for distinguishing among different forms of disrespect in terms of both where they occur and how severely they affect their victims. Such a doctrine can even recognize principled distinctions between disrespect and other immoralities, defining the common good (and thus the legitimate scope of political power) to encompass only other-regarding wrongs, even if some of those wrongs—such as the wrong of discrimination—must be explicated in terms of the moral characters of their perpetrators.

To be clear, the respect-based doctrine of political legitimacy developed in this chapter does not purport to provide a complete account of the legitimate scope of state power. The state, in other words, is not limited to prohibiting disrespectful conduct, but rather has authority to address a much wider range of wrongs as well. A respect-based doctrine of political legitimacy is thus meant to supplement, rather than supplant, other doctrines. Indeed, a respect-based doctrine of political legitimacy might even be compatible with the doctrines considered in Part I of this thesis, once purged of their antiperfectionist elements. We might, for example, invoke a perfectionist, respect-based doctrine of political legitimacy to justify antidiscrimination law's core prohibitions of direct discrimination and then turn to Rawls's conception of distributive justice (freed from the strictures of public reason) to explain why we must extend those prohibitions to indirect discrimination. Similarly, a Millian or Razian conception of harm might provide the most compelling account of the wrongfulness of many instances of discrimination, even if a respect-based doctrine of political legitimacy is required to reach the rest. To criticize the doctrines considered in Part I as incomplete is not to refute all their elements. But in order to justify the full extent of antidiscrimination law, the doctrines must ally themselves with some ancillary doctrine of political legitimacy, one that attends to discriminators' intentions. And they can form such an alliance only by jettisoning their antiperfectionism.

Moreover, while this thesis cannot canvass every possible perfectionist account of antidiscrimination law, it does seem that the most plausible account will resemble the one developed in this chapter and the previous one. For a perfectionist account will distinguish itself from its antiperfectionist rivals by attending to discriminators' intentions and explaining why they are so objectionable. Because such an explanation requires us to draw some connection between one's reasons for action and one's treatment of others, it

is difficult to see how we can provide the explanation without adverting to some cognitive notion like respect. Purely noncognitive notions such as distributive justice and harm—even when understood in perfectionist terms—seem insufficient for this task. Thus, in addition to arguing that the doctrines of political legitimacy considered in Part I must embrace perfectionism, this thesis contends that they must embrace something very close to the particular brand of perfectionism developed in Part II.

Conclusion

This thesis has said little about the actual practice of antidiscrimination law. Such an omission should come as little surprise, given the nature of the task this thesis set for itself. Whereas most of the actual practice of antidiscrimination law is concerned with *proving* the wrong of discrimination based on race or sex,¹ this thesis has been concerned with defining that wrong and assessing the legitimacy of governmental efforts to combat it. That is, this thesis has focused on relatively abstract questions of morality rather than more concrete questions of its institutionalization.² Consequently, insofar as this thesis has attended to the actual provisions of U.K. and U.S. antidiscrimination law, it has focused on those provisions that define the offense of discrimination, not those that articulate subsidiary doctrines for proving that offense. To be sure, the two sets of provisions are closely related: how we define the offense of discrimination partly determines how we go about proving it. But the tasks of definition and proof are nevertheless distinct, with the first involving issues of political theory and the second involving issues of legal practice and public policy.

This invocation of the distinction between morality and its institutionalization echoes contemporary liberalism's invocation of the same distinction to attempt to evade this thesis's polemical strategy, an issue considered in Chapter 1. It is important to recognize, however, that this thesis and contemporary liberalism draw the distinction in very different ways. As we saw in Chapter 1, contemporary liberals appear willing, at least in principle, to treat almost all aspects of antidiscrimination law—including the legal definition of discrimination itself—as matters of institutionalization. This thesis, by

¹ See, e.g., the body of case law expounding the burden-shifting framework for proving direct discrimination that was originally articulated by the U.S. Supreme Court in *McDonnell Douglas Corp. v. Green*, 411 U.S. 792 (1973).

² Admittedly, the morality/institutionalization distinction is vague, and this thesis has occasionally strayed into the grey area between the two concepts.

contrast, considers at least the legal definition of discrimination to be a question of first-order morality, a question of how individual moral agents should act. This thesis, in other words, supposes that the legal definition of discrimination tracks morality, even if the legal doctrines for proving that offense may take institutional considerations into account and thus depart from a more direct instantiation of moral principles. Neither demarcation of the morality/institutionalization distinction seems superior in the abstract. But as Chapter 5 attempted to show, we can derive a convincing account of the wrongfulness of discrimination from the legal definition of the conduct. Such an account, moreover, confirms our intuitions that discrimination, as currently defined by antidiscrimination law, is not just a legal offense, but also a moral wrong, and that antidiscrimination law is no ordinary piece of legislation, but an embodiment of some of our most fundamental moral convictions. These intuitions may very well be mistaken, but they seem to be firmly rooted—and they are ones that almost all contemporary liberals profess to share.

Once we turn from morality to its institutionalization, from defining discrimination to proving it, we realize that the wrong of discrimination, as defined in Chapter 5, is extremely difficult to prove. Recall that discrimination is wrong because it is wrong to take another person's race or sex as a reason to treat her less favorably than one would treat people of other races or the other sex. Acting on such an intention precludes having an appropriately respectful attitude toward the victim, one that would regard her as possessing ultimate value. Because the wrongfulness of discrimination thus inheres in the intentions of discriminators, we must scrutinize a person's intentions in order to prove that he discriminated. Obviously, the law cannot read people's minds. Therefore, barring some sort of confession by the discriminator, the law will rarely be able to prove definitively that an alleged discriminator acted with a discriminatory intention. Instead, the law will often have to rely on circumstantial evidence of the discriminator's

intentions, such as the discriminator's treatment of the victim's relevant "comparators"³ or aggregate statistics summarizing the discriminator's treatment of people of different races and sexes.⁴ Such circumstantial evidence is, of course, an imperfect means of proving alleged discriminators' intentions, one that will occasionally condemn the innocent and exonerate the guilty.⁵ While the legal definition of discrimination may track morality, the practice of antidiscrimination law often will not.

At least in the U.S. and U.K. contexts, this divergence between the morality of discrimination and the practice of antidiscrimination law might not be particularly disconcerting, for neither country criminalizes discrimination.⁶ Rather, as we saw in Chapter 3, both countries seem to model antidiscrimination law on tort law, requiring discriminators to compensate their victims but declining to punish discriminators for their wrongdoing. Because civil compensation, unlike criminal punishment, does not express blame,⁷ we might think that antidiscrimination law can tolerate a significant number of false positives, that it rightly forgoes the higher burdens of proof and other pro-defendant procedural safeguards of criminal law for the more pro-plaintiff procedures of tort law.

But in fact, discrimination differs from many (though by no means all) torts in that its wrongfulness is defined in terms of the moral characters of its perpetrators. For to conclude that someone has discriminated is to conclude that he has been disrespectful, that he has failed to cultivate appropriate attitudes toward people of a particular race or sex. Perhaps antidiscrimination law is nevertheless right to deny discrimination

³ See, e.g., *Ash v. Tyson Foods, Inc.*, 546 U.S. 454 (2006).

⁴ See, e.g., *Teamsters v. United States*, 431 U.S. 324 (1977).

⁵ Antidiscrimination law is thus an example of what John Rawls calls "imperfect procedural justice" (John Rawls, *A Theory of Justice*, rev. ed. (Cambridge, Mass.: Harvard University Press, 1999), 74-75).

⁶ Some other countries, however, do criminalize discrimination. See generally Julie C. Suk, "Procedural Path Dependence: Discrimination and the Civil-Criminal Divide," *Washington University Law Review* 85 (2008): 1315-71.

⁷ See Joel Feinberg, *Doing and Deserving* (Princeton: Princeton University Press, 1970), ch. 5.

defendants the heightened procedural safeguards of the criminal law;⁸ after all, discrimination defendants do not face punishment for their wrongdoing. But despite there being no prospect of punishment, the connection between the wrongfulness of discrimination and the characters of discriminators ensures that a charge of discrimination will carry a connotation of condemnation.

At least through its prohibitions against direct discrimination, then, antidiscrimination law censures discriminators in addition to compensating victims, an expressive dimension that reflects ordinary moral practice, from the morally laden connotations evoked by the term “discrimination” to the contempt we tend to show discriminators. Do these moralistic features of antidiscrimination law help to improve the condition of racial minorities and women? Perhaps not. Perhaps they distract us from more pressing concerns of distributive justice, concerns antidiscrimination law might be able to address more effectively if it instead sought to structure workplaces and other public venues to minimize bias against, and maximize opportunities for, racial minorities and women.⁹ But this is not the strategy that contemporary antidiscrimination law pursues. So long as we remain committed to antidiscrimination law in its current form, we must attend to discriminators’ characters. And to attend to discriminators’ characters, we must be willing to espouse perfectionism.

⁸ In practice, of course, judges and other legal officials can (and often do) bend ostensibly pro-plaintiff procedures in pro-defendant directions.

⁹ For examples of this structuralist approach, see Tristin K. Green, “Work Culture and Discrimination,” *California Law Review* 93 (2005): 623-84; and Susan Sturm, “Second Generation Employment Discrimination: A Structural Approach,” *Columbia Law Review* 101 (2001): 458-568.

Works Cited

BOOKS AND ARTICLES

- Alexander, Larry. "Rules, Rights, Options, and Time." *Legal Theory* 6 (2000): 391-404.
- . "What Is Freedom of Association and What Is Its Denial?" *Social Philosophy and Policy* 25 (2008): 1-21.
- . "What Makes Wrongful Discrimination Wrong? Biases, Preferences, Stereotypes, and Proxies." *University of Pennsylvania Law Review* 141 (1992): 149-219.
- Alexander, Larry and Kevin Cole. "Discrimination by Proxy." *Constitutional Commentary* 14 (1997): 453-63.
- Anderson, Elizabeth. *The Imperative of Integration*. Princeton: Princeton University Press, 2010.
- Anscombe, G.E.M. *Intention*. 2nd ed. Oxford: Basil Blackwell, 1963.
- Aristotle. *Nicomachean Ethics*. Translated by Christopher Rowe. Oxford: Oxford University Press, 2002.
- Arneson, Richard J. "Against Rawlsian Equality of Opportunity." *Philosophical Studies* 93 (1999): 77-112.
- . "Perfectionism and Politics." *Ethics* 111 (2000): 37-63.
- . "Sophisticated Rule Consequentialism: Some Simple Objections." *Philosophical Issues* 15 (2005): 235-51.
- . "What Is Wrongful Discrimination?" *San Diego Law Review* 43 (2006): 775-808.
- Barry, Brian M. *Culture and Equality: An Egalitarian Critique of Multiculturalism*. Cambridge: Polity Press, 2001.
- . *Justice as Impartiality*. Oxford: Oxford University Press, 1996.
- Berlin, Isaiah. *Four Essays on Liberty*. Oxford: Oxford University Press, 1990.
- Bird, Colin. "Harm Versus Sovereignty: A Reply to Ripstein." *Philosophy & Public Affairs* 35 (2007): 179-94.
- Brest, Paul. "The Supreme Court 1975 Term—Foreword: In Defense of the Antidiscrimination Principle." *Harvard Law Review* 90 (1976): 1-55.
- Buchanan, Allen E. *Justice, Legitimacy, and Self-determination: Moral Foundations for International Law*. Oxford: Oxford University Press, 2004.

- . "Political Legitimacy and Democracy." *Ethics* 112 (2002): 689-719.
- Casal, Paula. "Why Sufficiency Is Not Enough." *Ethics* 117 (2007): 296-326.
- Cavanagh, Matt. *Against Equality of Opportunity*. Oxford: Oxford University Press, 2003.
- Chan, Joseph. "Legitimacy, Unanimity, and Perfectionism." *Philosophy & Public Affairs* 29 (2000): 5-42.
- Cohen, G.A. "Incentives, Inequality, and Community." In *The Tanner Lectures on Human Values*, Vol. 13, edited by Grethe Peterson. Salt Lake City: University of Utah Press, 1992.
- . "Where the Action Is: On the Site of Distributive Justice." *Philosophy & Public Affairs* 26 (1997): 3-30.
- Cohen, Joshua. "A More Democratic Liberalism." *Michigan Law Review* 92 (1994): 1503-46.
- Coleman, Jules. *Risks and Wrongs*. Cambridge: Cambridge University Press, 1992.
- Cowen, Tyler. "Rule Consequentialism Makes Sense After All." *Social Philosophy and Policy* 28 (2011): 212-31.
- Crisp, Roger. "Modern Moral Philosophy and the Virtues." In *How Should One Live? Essays on the Virtues*, edited by Roger Crisp. Oxford: Oxford University Press, 1996.
- Daniels, Norman. "Democratic Equality: Rawls's Complex Egalitarianism." In *The Cambridge Companion to Rawls*, edited by Samuel Freeman. Cambridge: Cambridge University Press, 2003.
- Darwall, Stephen L. "Two Kinds of Respect." *Ethics* 88 (1977): 36-49.
- Davis, Nancy (Ann). "Contemporary Deontology." In *A Companion to Ethics*, edited by Peter Singer. Oxford: Blackwell, 1991.
- Duff, R.A. "Harms and Wrongs." *Buffalo Criminal Law Review* 5 (2001): 13-45.
- Dworkin, Ronald. *Law's Empire*. Cambridge, Mass.: Harvard University Press, 1986.
- . *A Matter of Principle*. Cambridge, Mass.: Harvard University Press, 1985.
- . *Sovereign Virtue: The Theory and Practice of Equality*. Cambridge, Mass.: Harvard University Press, 2000.
- Ely, John Hart. *Democracy and Distrust: A Theory of Judicial Review*. Cambridge, Mass.: Harvard University Press, 1980.

- Epstein, Richard A. *Forbidden Grounds: The Case Against Employment Discrimination Laws*. Cambridge, Mass.: Harvard University Press, 1992.
- . "The Harm Principle—And How It Grew." *University of Toronto Law Journal* 45 (1995): 369-417.
- Estlund, Cynthia. *Working Together: How Workplace Bonds Strengthen a Diverse Democracy*. Oxford: Oxford University Press, 2003.
- Feinberg, Joel. *Doing and Deserving: Essays in the Theory of Responsibility*. Princeton: Princeton University Press, 1970.
- . *Freedom and Fulfillment: Philosophical Essays*. Princeton: Princeton University Press, 1992.
- . *Harm to Others*. Vol. 1 of *The Moral Limits of the Criminal Law*. Oxford: Oxford University Press, 1984.
- . *Harm to Self*. Vol. 3 of *The Moral Limits of the Criminal Law*. Oxford: Oxford University Press, 1986.
- . *Harmless Wrongdoing*. Vol. 4 of *The Moral Limits of the Criminal Law*. Oxford: Oxford University Press, 1990.
- Finnis, John. *Aquinas*. Oxford: Oxford University Press, 1998.
- . *Intention and Identity*. Vol. 2 of *Collected Essays*. Oxford: Oxford University Press, 2011.
- . "Intention in Tort Law." In *Philosophical Foundations of Tort Law*, edited by David G. Owen. Oxford: Oxford University Press, 1995.
- . "Is Natural Law Theory Compatible with Limited Government?" In *Natural Law, Liberalism, and Morality: Contemporary Essays*, edited by Robert P. George. Oxford: Oxford University Press, 1996.
- . "Legal Enforcement of 'Duties to Oneself': Kant v. Neo-Kantians." *Columbia Law Review* 87 (1987): 433-56.
- . *Natural Law and Natural Rights*. Oxford: Oxford University Press, 1980.
- Fiss, Owen. "Groups and the Equal Protection Clause." *Philosophy & Public Affairs* 5 (1976): 107-77.
- Frankfurt, Harry G. *The Importance of What We Care About*. Cambridge: Cambridge University Press, 1988.
- . *Necessity, Volition, and Love*. Cambridge: Cambridge University Press, 1999.

- Freeman, Alan David. "Legitimizing Racial Discrimination Through Antidiscrimination Law: A Critical Review of Supreme Court Doctrine." *Minnesota Law Review* 62 (1978): 1049-1119.
- Freeman, Samuel. "Illiberal Libertarians: Why Libertarianism Is Not a Liberal View." *Philosophy & Public Affairs* 30 (2002): 105-51.
- Fuller, Lon L. *The Morality of Law*. Rev. ed. New Haven, Conn.: Yale University Press, 1969.
- Garcia, J.L.A. "Current Conceptions of Racism: A Critical Examination of Some Recent Social Philosophy." *Journal of Social Philosophy* 28 (1997): 5-42.
- . "The Heart of Racism." In *Race and Racism*, edited by Bernard Boxill. Oxford: Oxford University Press, 2001.
- . "Philosophical Analysis and the Moral Concept of Racism." *Philosophy & Social Criticism* 25 (1999): 1-32.
- Gardner, John. "Justifications and Reasons." In *Harm and Culpability*, edited by A.P. Simester and A.T.H. Smith. Oxford: Oxford University Press, 1996.
- . "Liberals and Unlawful Discrimination." *Oxford Journal of Legal Studies* 9 (1989): 1-22.
- . "On the Ground of Her Sex(uality)." *Oxford Journal of Legal Studies* 18 (1998): 167-87.
- . "Private Activities and Personal Autonomy: At the Margins of Anti-Discrimination Law." In *Discrimination: The Limits of Law*, edited by Bob Hepple and Erica M. Szyszczak. London: Mensell, 1992.
- . "Wrongs and Faults." In *Appraising Strict Liability*, edited by A.P. Simester. Oxford: Oxford University Press, 2005.
- Gardner, John and Stephen Shute. "The Wrongness of Rape." In *Oxford Essays in Jurisprudence: Fourth Series*, edited by Jeremy Horder. Oxford: Oxford University Press, 2000.
- George, Robert P. *Making Men Moral: Civil Liberties and Public Morality*. Oxford: Oxford University Press, 1993.
- Glasgow, Joshua. "Racism as Disrespect." *Ethics* 120 (2009): 64-93.
- Goffman, Erving. *Stigma: Notes on the Management of Spoiled Identity*. New York: Simon & Schuster, Inc., 1963.
- Goldman, Alan H. "Affirmative Action." *Philosophy & Public Affairs* 5 (1976): 178-95.
- . *Justice and Reverse Discrimination*. Princeton: Princeton University Press, 1979.

- Green, Leslie. *The Authority of the State*. Oxford: Oxford University Press, 1988.
- . "Two Worries About Respect for Persons." *Ethics* 120 (2010): 212-31.
- Green, Tristin K. "Work Culture and Discrimination." *California Law Review* 93 (2005): 623-84.
- Greenawalt, Kent. "Freedom of Association and Religious Association." In *Freedom of Association*, edited by Amy Gutmann. Princeton: Princeton University Press, 1998.
- Griffin, James. "Discrepancies Between the Best Philosophical Account of Human Rights and the International Law of Human Rights." *Proceedings of the Aristotelian Society* 101 (2001): 1-28.
- . "First Steps in an Account of Human Rights." *European Journal of Philosophy* 9 (2001): 306-27.
- Gutmann, Amy. "Freedom of Association: An Introductory Essay." In *Freedom of Association*, edited by Amy Gutmann. Princeton: Princeton University Press, 1998.
- . "Responding to Racial Injustice." In K. Anthony Appiah and Amy Gutmann, *Color Conscious: The Political Morality of Race*. Princeton: Princeton University Press, 1996.
- Harcourt, Bernard E. "The Collapse of the Harm Principle." *Journal of Criminal Law & Criminology* 90 (1999): 109-94.
- Hart, H.L.A. *Law, Liberty, and Morality*. Stanford: Stanford University Press, 1963.
- Hellman, Deborah. *When Is Discrimination Wrong?* Cambridge, Mass.: Harvard University Press, 2008.
- Hill, Thomas E., Jr. *Dignity and Practical Reason in Kant's Moral Theory*. Ithaca: Cornell University Press, 1992.
- . *Respect, Pluralism, and Justice: Kantian Perspectives*. Oxford: Oxford University Press, 2000.
- Holmes, Elisa. "Anti-Discrimination Rights Without Equality." *Modern Law Review* 68 (2005): 175-94.
- Hooker, Brad. "Rule-Consequentialism." In *The Blackwell Guide to Ethical Theory*, edited by Hugh LaFollette. Oxford: Blackwell, 2000.
- Hurka, Thomas. "Indirect Perfectionism: Kymlicka on Liberal Neutrality." *Journal of Political Philosophy* 3 (1995): 36-57.

- . *Perfectionism*. Oxford: Oxford University Press, 1996.
- Kang, Jerry and Kristin Lane. "Seeing Through Color Blindness: Implicit Bias and the Law." *UCLA Law Review* 58 (2010): 465-520.
- Kant, Immanuel. *Groundwork of the Metaphysics of Morals*. Translated and edited by Mary Gregor. Cambridge: Cambridge University Press, 1997.
- . *The Metaphysics of Morals*. Translated and edited by Mary Gregor. Cambridge: Cambridge University Press, 1996.
- Kelly, Erin and Lionel McPherson. "On Tolerating the Unreasonable." *Journal of Political Philosophy* 9 (2001): 38-55.
- Knowles, Dudley R. "A Reformulation of the Harm Principle." *Political Theory* 6 (1978): 233-46.
- Koppelman, Andrew. *Antidiscrimination Law and Social Equality*. New Haven: Yale University Press, 1996.
- Kordana, Kevin A. and David H. Blankfein Tabachnick. "The Rawlsian View of Private Ordering." *Social Philosophy and Policy* 25 (2008): 288-307.
- Korsgaard, Christine M. *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press, 1996.
- Lamond, Grant. "Coercion and the Nature of Law." *Legal Theory* 7 (2001): 35-57.
- Larmore, Charles E. *Patterns of Moral Complexity*. Cambridge: Cambridge University Press, 1987.
- Lawrence Charles R., III. "The Id, the Ego, and Equal Protection: Reckoning with Unconscious Racism." *Stanford Law Review* 39 (1987): 317-88.
- Lippert-Rasmussen, Kasper. "The Badness of Discrimination." *Ethical Theory and Moral Practice* 9 (2006): 167-85.
- Locke, John. *Two Treatises of Government*. Edited by Peter Laslett. Cambridge: Cambridge University Press, 1998.
- Loury, Glenn C. *The Anatomy of Racial Inequality*. Cambridge, Mass.: Harvard University Press, 2003.
- McCabe, David. "Joseph Raz and the Contextual Argument for Liberal Perfectionism." *Ethics* 111 (2001): 493-522.
- McCrudden, Christopher. "Institutional Discrimination." *Oxford Journal of Legal Studies* 2 (1982): 303-67.

- Meyer, Lukas H. "Past and Future: The Case for a Threshold Notion of Harm." In *Rights, Culture, and the Law: Themes from the Legal and Political Philosophy of Joseph Raz*, edited by Lukas H. Meyer, Stanley L. Paulson, and Thomas W. Pogge. Oxford: Oxford University Press, 2003.
- Mill, John Stuart. *On Liberty and Other Essays*. Edited by John Gray. Oxford: Oxford University Press, 1998.
- Moreau, Sophia. "What Is Discrimination?" *Philosophy & Public Affairs* 38 (2010): 143-79.
- Mulhall, Stephen and Adam Swift. *Liberals and Communitarians*. 2nd ed. Oxford: Blackwell Publishing, 1996.
- Murphy, Liam B. "Institutions and the Demands of Justice." *Philosophy & Public Affairs* 27 (1998): 251-91.
- Nagel, Thomas. "Equal Treatment and Compensatory Discrimination." *Philosophy & Public Affairs* 2 (1973): 348-63.
- . *Equality and Partiality*. Oxford: Oxford University Press, 1991.
- . "John Rawls and Affirmative Action." *Journal of Blacks in Higher Education* 39 (2003): 82-84.
- . "Justice and Nature." *Oxford Journal of Legal Studies* 17 (1997): 303-21.
- . "Moral Conflict and Political Legitimacy." *Philosophy & Public Affairs* 16 (1987): 215-40.
- . "Rawls and Liberalism." In *The Cambridge Companion to Rawls*, edited by Samuel Freeman. Cambridge: Cambridge University Press, 2003.
- Nussbaum, Martha C. "Perfectionist Liberalism and Political Liberalism." *Philosophy & Public Affairs* 39 (2011): 3-45.
- Okin, Susan Moller. *Justice, Gender, and the Family*. New York: Basic Books, 1989.
- O'Neill, Onora. *Constructions of Reason*. Cambridge: Cambridge University Press, 1989.
- Patten, Alan. "Liberal Neutrality: A Reinterpretation and Defense." *Journal of Political Philosophy* (forthcoming 2011).
- Pettit, Philip. "Agency-Freedom and Option-Freedom." *Journal of Theoretical Politics* 15 (2003): 387-403.
- . "Consequentialism." In *A Companion to Ethics*, edited by Peter Singer. Oxford: Blackwell, 1991.

- . *Republicanism: A Theory of Freedom and Government*. Oxford: Oxford University Press, 1997.
- . *A Theory of Freedom: From the Psychology to the Politics of Agency*. Oxford: Oxford University Press, 2001.
- Pogge, Thomas W. *Realizing Rawls*. Ithaca: Cornell University Press, 1989.
- Posner, Richard A. *The Economics of Justice*. Cambridge, Mass.: Harvard University Press, 1981.
- Putnam, Robert. *Bowling Alone: The Collapse and Revival of American Community*. New York: Simon and Schuster, 2000.
- Quong, Jonathan. *Liberalism Without Perfection*. Oxford: Oxford University Press, 2011.
- . "The Rights of Unreasonable Citizens." *Journal of Political Philosophy* 12 (2004): 314-35.
- Rawls, John. *Justice as Fairness: A Restatement*. Edited by Erin Kelly. Cambridge, Mass.: Harvard University Press, 2001.
- . *The Law of Peoples, with "The Idea of Public Reason Revisited."* Cambridge, Mass.: Harvard University Press, 1999.
- . *Political Liberalism*. Paperback ed. New York: Columbia University Press, 1996.
- . *A Theory of Justice*. Rev. ed. Cambridge, Mass.: Harvard University Press, 1999.
- . "Two Concepts of Rules." *Philosophical Review* 16 (1955): 3-32.
- Raz, Joseph. *The Authority of Law: Essays on Law and Morality*. Oxford: Oxford University Press, 1979.
- . *Between Authority and Interpretation: On the Theory of Law and Practical Reason*. Oxford: Oxford University Press, 2009.
- . *Engaging Reason*. Oxford: Oxford University Press, 2002.
- . *Ethics in the Public Domain: Essays in the Morality of Law and Politics*. Rev. ed. Oxford: Oxford University Press, 2001.
- . "Facing up: A Reply." *Southern California Law Review* 62 (1989): 1153-1236.
- . "Liberty and Trust." In *Natural Law, Liberalism, and Morality*, edited by Robert P. George. Oxford: Oxford University Press, 1996.
- . *The Morality of Freedom*. Oxford: Oxford University Press, 1986.

- . "On Frankfurt's Explanation of Respect for People." In *Contours of Agency: Essays on Themes from Harry Frankfurt*, edited by Sarah Buss and Lee Overton. Cambridge, Mass.: MIT Press, 2002.
- . "On Respect, Authority, and Neutrality: A Response." *Ethics* 120 (2010): 279-301.
- . "On the Guise of the Good." In *Desire, Practical Reason, and the Good*, edited by Sergio Tenenbaum. Oxford: Oxford University Press, 2010.
- . *Practical Reason and Norms*. 2nd ed. Oxford: Oxford University Press, 1999.
- . *Value, Respect, and Attachment*. Cambridge: Cambridge University Press, 2001.
- Réaume, Denise G. "Discrimination and Dignity." *Louisiana Law Review* 63 (2003): 645-95.
- . "Harm and Fault in Discrimination Law: The Transition from Intentional to Adverse Effect Discrimination." *Theoretical Inquiries in Law* 2 (2001): 349-386.
- Richards, David A.J. *Sex, Drugs, Death, and the Law: An Essay on Human Rights and Overcriminalization*. Totowa, N.J.: Rowman & Littlefield, 1982.
- . *Toleration and the Constitution*. Oxford: Oxford University Press, 1986.
- Ripstein, Arthur. "Beyond the Harm Principle." *Philosophy & Public Affairs* 34: (2006): 215-45.
- . *Equality, Responsibility, and the Law*. Cambridge: Cambridge University Press, 1999.
- . *Force and Freedom: Kant's Legal and Political Philosophy*. Cambridge, Mass.: Harvard University Press, 2009.
- . "Legal Moralism and the Harm Principle: A Rejoinder." *Philosophy & Public Affairs* 35 (2007): 195-201.
- Rosenblum, Nancy L. "Compelled Association: Public Standing, Self-Respect, and the Dynamic of Exclusion." In *Freedom of Association*, edited by Amy Gutmann. Princeton: Princeton University Press, 1998.
- Ross, David. *The Right and the Good*. Edited by Philip Stratton-Lake. Oxford: Oxford University Press, 2002.
- Sadurski, Wojciech. "Joseph Raz on Liberal Neutrality and the Harm Principle." *Oxford Journal of Legal Studies* 10 (1990): 122-33.
- Scanlon, T.M. *The Difficulty of Tolerance: Essays in Political Philosophy*. Cambridge: Cambridge University Press, 2003.
- . *Moral Dimensions: Permissibility, Meaning, Blame*. Cambridge, Mass.: Harvard University Press, 2008.

- Scheffler, Samuel. "What Is Egalitarianism?" *Philosophy & Public Affairs* 31 (2003): 5-39.
- Shelby, Tommie. "Is Racism in the 'Heart'?" *Journal of Social Philosophy* 33 (2002): 411-20.
- . "Justice, Deviance, and the Dark Ghetto." *Philosophy & Public Affairs* 35 (2007): 126-60.
- . "Race and Social Justice: Rawlsian Considerations." *Fordham Law Review* 72 (2004): 1697-1714.
- Sher, George. *Beyond Neutrality: Perfectionism and Politics*. Cambridge: Cambridge University Press, 1997.
- . *Who Knew? Responsibility Without Awareness*. Oxford: Oxford University Press, 2009.
- Shiffrin, Seana Valentine. "Race, Labor, and the Fair Equality of Opportunity Principle." *Fordham Law Review* 72 (2004): 1643-75.
- . "What Is Really Wrong with Compelled Association?" *Northwestern University Law Review* 99 (1995): 839-88.
- Shin, Patrick S. "The Substantive Principle of Equal Treatment." *Legal Theory* 15 (2009): 149-72.
- Shultz, Vicki. "Life's Work." *Columbia Law Review* 100 (2000): 1881-1964.
- Simmons, A. John. "Justification and Legitimacy." *Ethics* 109 (1999): 739-71.
- Stanton-Ife, John. "The Limits of Law." *Stanford Encyclopedia of Philosophy*, <http://plato.stanford.edu/entries/law-limits/>.
- Stewart, Hamish. "Harms, Wrongs, and Set-Backs in Feinberg's *Moral Limits of the Criminal Law*." *Buffalo Criminal Law Review* 5 (2001): 47-67.
- . "The Limits of the Harm Principle." *Criminal Law and Philosophy* 4 (2010): 17-35.
- Sturm, Susan. "Second Generation Employment Discrimination: A Structural Approach." *Columbia Law Review* 101 (2001): 458-568.
- Suk, Julie C. "Procedural Path Dependence: Discrimination and the Civil-Criminal Divide." *Washington University Law Review* 85 (2008): 1315-71.
- Sunstein, Cass R. "The Anticaste Principle." *Michigan Law Review* 92 (1994): 2410-55.

- . "Should Sex Equality Law Apply to Religious Institutions?" In *Is Multiculturalism Bad for Women?*, edited by Joshua Cohen, Matthew Howard, and Martha C. Nussbaum. Princeton: Princeton University Press, 1999.
- Sverdlik, Steven. *Motive and Rightness*. Oxford: Oxford University Press, 2011.
- . "Motive and Rightness." *Ethics* 106 (1996): 327-49.
- Tadros, Victor. "Harm, Sovereignty, and Prohibition." *Legal Theory* 17 (2011): 35-65.
- . "Independence Without Interests?" *Oxford Journal of Legal Studies* 31 (2011): 193-213.
- Tasioulas, John. "Human Rights, Universality and the Values of Personhood: Retracing Griffin's Steps." *European Journal of Philosophy* 10 (2002): 79-100.
- Taylor, Robert S. "Hate Speech, the Priority of Liberty, and the Temptations of Nonideal Theory." *Ethical Theory and Moral Practice* (forthcoming 2011).
- . "Rawls's Defense of the Priority of Liberty: A Kantian Reconstruction." *Philosophy & Public Affairs* 31 (2003): 246-71.
- Thomson, Judith Jarvis. "Physician-Assisted Suicide: Two Arguments." *Ethics* 109 (1999): 497-518.
- . "Preferential Hiring." *Philosophy & Public Affairs* 2 (1973): 364-84.
- . "Self-Defense." *Philosophy & Public Affairs* 20 (1991): 283-310.
- Volokh, Eugene. "What Speech Does 'Hostile Work Environment' Harassment Law Restrict?" *Georgetown Law Journal* 85 (1997): 627-48.
- Waldron, Jeremy. "Autonomy and Perfectionism in Raz's *Morality of Freedom*." *Southern California Law Review* 62 (1989): 1097-1152.
- . *Liberal Rights: Collected Papers, 1981-1991*. Cambridge: Cambridge University Press, 1993.
- . "Moral Autonomy and Personal Autonomy." In *Autonomy and the Challenges to Liberalism: New Essays*, edited by John Christman and Joel Anderson. Cambridge: Cambridge University Press, 2005.
- Wasserstrom, Richard. "Racism, Sexism, and Preferential Treatment." *UCLA Law Review* 24 (1976):
- Wertheimer, Alan. "Jobs, Qualifications, and Preferences." *Ethics* 94 (1983): 581-622.
- . "Reflections on Discrimination." *San Diego Law Review* 43 (2006): 945-80.

- White, Stuart. "Freedom of Association and the Right to Exclude." *Journal of Political Philosophy* 5 (1997): 373-91.
- Williams, Bernard. *Moral Luck*. Cambridge: Cambridge University Press, 1981.
- Wintemute, Robert. *Sexual Orientation and Human Rights: The United States Constitution, the European Convention, and the Canadian Charter*. Oxford: Oxford University Press, 1995.
- Wolff, Robert Paul. *In Defense of Anarchism*. Berkeley, Calif.: University of California Press, 1998.
- Young, Iris. *Justice and the Politics of Difference*. Princeton: Princeton University Press, 1990.
- Yuracko, Kimberly A. *Perfectionism and Contemporary Feminist Values*. Bloomington, IN: Indiana University Press, 2003.
- . "Private Nurses and Playboy Bunnies: Explaining Permissible Sex Discrimination." *California Law Review* 92 (2004): 147-213.
- . "Sameness, Subordination, and Perfectionism: Toward a More Complete Theory of Employment Discrimination Law." *San Diego Law Review* 43 (2006): 857-97.

STATUTES

- U.K. Equality Act 2010.
- U.K. Race Relations Act 1976.
- U.K. Sex Discrimination Act 1975.
- U.S. Civil Rights Act of 1964.

CASES

- James v. Eastleigh Borough Council*, [1990] 2 AC 751.
- Ash v. Tyson Foods, Inc.*, 546 U.S. 454 (2006).
- Desert Palace, Inc. v. Costa*, 539 U.S. 90 (2003).
- Griggs v. Duke Power Co.*, 401 U.S. 424 (1971).
- McDonnell Douglas Corp. v. Green*, 411 U.S. 792 (1973).
- Teamsters v. United States*, 431 U.S. 324 (1977).